University of Alberta

Taxonomies, Knowledge, and Artifacts; Interactivity in Category Learning

by

Michael Hoff

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Psychology

Spring 2010
Edmonton, Alberta

Examining Committee

Thomas Spalding, Psychology

Christina Gagné, Psychology

Peter Dixon, Psychology

Patrick Bolger, Linguistics

Scott Allen, Psychology, University of Lethbridge

ABSTRACT

Many adult concepts can be represented in taxonomies – hierarchical systems in which concepts are differentiated into varying levels of abstraction (e.g., musical instrument, wind instrument, flute) related by class inclusion (a flute is a wind instrument and a wind instrument is a musical instrument). Indeed, most natural kinds (e.g., whale, tree) and artifacts (e.g., flute, truck) are generally believed to fall within taxonomies. Moreover, in real world contexts, concepts are probably rarely learned as explicitly contrasting sets existing completely outside of known taxonomies (that is, one might not learn cats vs. dogs without also learning that both are types of animals, and that both include more specific subcategories). Surprisingly, relatively little research has been done on the learning of categories that are hierarchically structured. The present study began an investigation into how adults learned new concepts that are hierarchically structured. In Experiment 1, participants learned to classify items at one taxonomic level then at a later time classified items at either the same or a different level. The results suggested that people were unable to clearly detect the relationship among alternate levels of the hierarchy prior to exposure of those levels. However, results in Experiment 1 also suggested that learning multiple categories might lead to deeper understanding of how features transfer or generalize to higher taxonomic levels. The remaining experiments addressed more explicitly the influence of hierarchical structures on category learning by including prototype and control items, along with artificial and knowledge-based category labels. Results from these experiments indicated that, at least within the parameters of this study, prior experience cued by knowledge-based category labels interacted adversely with abstract materials and interfered with mapping of item information to categories.

Moreover, when the relationship between categories and item information is unclear, generalization might be one important means by which people categorize.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisor Tom Spalding for his keen insights, endurance, and patience. From beginning to end this project has been one of collaboration and Tom was there at every step.

Thank you to Christina Gagne′ for her comments, time and effort but most importantly for her encouragement.

Thank you to Peter Dixon for his exhaustive and helpful comments, all of which went far beyond the call of duty.

Special thanks to Scott Allen and Patrick Bolger for joining this circus; to Ken Prkachin, Glenda Prkachin, Milly Saville, Craig Norton, Jim Kinnear, Bernd Kohler, Bob Hacket, Toby Hacket, Michael Pratt, Lauren Figueredo and Hongbo Ji (no better office mate).

Finally, I would like to thank my wife Shari for her support, understanding and allowing me the freedom to pursue my goals.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Taxonomies, Knowledge, and Artifacts; Interactivity in Category Learning

Overview

Partitioning the world into meaningful categories is a formidable task, especially considering the vast number of comparisons that must be made during the categorization process. For instance, it is likely that when we first learn the category *dog* our comparison extends beyond other types of dogs; indeed potentially any object of natural kind experienced (or being experienced) by the agent is a candidate for comparison. Furthermore, the specific features of the object, its taxonomic position, the prior experience of the agent, and the context in which the decision occurs influence the categorization process. Given the sheer number of variables influencing a category decision the question of how the experimenter is able to capture the learning experience is very real.

Traditionally cognitive psychology has approached this task by carving the universe into distinct domains. Indeed a number of studies have demonstrated the effect of domain specific knowledge in categorization (Murphy & Medin, 1985; Pazzani, 1991; Wattenmaker, 1995). As a result, some have suggested (e.g., Hirschfeld & Gelman, 1994) that a categorization approach spanning multiple domains (e.g., artifacts, real world natural kinds, taxonomy, social categories) is problematic. However, it can be argued (see e.g., Bruner, Goodnow, & Austin, 1956) that a failure to merge domains results in only a limited understanding of how different domains interact to influence the category learning process.

One concern of this proposal is how various domains interact to influence category learning. For example, many studies exploring the effect of taxonomy on learning have focused narrowly on issues of either privileged taxonomic levels (see e.g., Rosch, 1975,1976; Murphy & Smith, 1982), psychological reality of taxonomic structures (see e.g., Sloman, 1993, 1998), or age of acquisition for taxonomic structures (see e.g., Markman & Callanan, 1984). Similarly, investigations of how knowledge influences category learning often occur under conditions where category structures already support coarse knowledge representations (see e.g., Kaplan & Murphy, 2000; Murphy & Allopenna, 1994; Spalding & Murphy, 1996; Ross & Murphy, 1999). Rarely (at least in the adult literature) have the influences of knowledge and taxonomy been explored within the context of structures having very weak connections to prior knowledge. The importance of using such structures is apparent when considering conceptual development. The use of fully formed or partially formed structures fails to capture initial learning experiences. This proposal introduces a new paradigm that is intended to capture such experiences, thereby exploring traditional questions such as whether there are natural tendencies toward using taxonomic structures. Additionally, new questions are addressed; for example, how does learning abstract item structures interact with prior knowledge within taxonomic structures?

Introduction

Categorization and concept are terms that cognitive psychologists use to describe how people conceptualize and understand the world around them. Although there is some general agreement that categorization is a fundamental cognitive process, and concepts are mental constructs, it is hard to find a consensual view on what the terms *category* and *concept* really represent.

According to Malt (1995), psychologists have failed to reach a consensus on the relative contribution of the environment versus the categorizer in determining categories. While some hold that the environment is highly structured and that the categorizer forms categories by directly recognizing structure in the world (Rosch, 1978), others hold that category formation is heavily influenced by cognitive processes that direct how the world is perceived (Murphy & Medin, 1985). Hampton and Dubois (1993) have suggested that controversies surrounding formalization and application of the terms *category* and *concept* result from cognitive science being a relatively young discipline where rival theories of categorization have led to a high degree of disparity between terms. One thing that will become evident as the major theories of categorization are surveyed below is that this is a diverse field that has undergone rapid paradigm shifts.

Since the establishment of cognitive science in the 1950's, several major psychological theories of classification have been advanced. The present proposal will focus on four of these theories: the classical view, the prototype theory, the exemplar theory, and the knowledge view.

1.1 Theories of Categorization

1.1.1 Classical View

The classical theory, which is one of the longest held views of categorization (Murphy, 2002), holds that the world displays a universal taxonomic order. All natural kinds belong to classes that are related by type relations and form a hierarchical structure. Concepts are considered terms, which consist of defining features that constitute a condition for category membership. Additionally, all classes are characterized by a membership condition that all objects must meet in order to be a member of the class. A concept represents a set of defining features that are independently necessary and jointly sufficient for category membership. For example, a triangle is a closed geometric figure having 3 sides, and interior angles that sum to 180 degrees. For a figure to meet the necessary definition of a triangle it must contain these defining characteristics otherwise it cannot be a triangle. So if a figure does not have interior angles that sum to 180 degrees it is not a triangle. The sufficient condition for a figure's classification as a triangle is met if it has all the features that define it as a triangle (e.g., not only has interior angles that sum to 180 degrees but a closed geometric figure having 3 sides).

Hull (1920) provides one of the earliest examples of experimental research on concept learning (Murphy, 2002) and was in large part influenced by the classical view. In his experiment Hull used artificial category stimuli. Because artificial stimuli are abstract and meaningless they are often used in order to control for affects of prior knowledge, learning, etc. Hull created twelve categories of twelve stimuli each. The stimuli were in the form of distorted Chinese characters. A meaningless sound was associated with each of the categories (i.e., "oo", "yer") and represented the category name. The categorization task involved participants responding phonetically to each of the stimuli. Consistent with the classical view, response accuracy was dependent on detecting the necessary and sufficient characteristics imbedded in each of the stimuli. For example, any stimuli containing a check-mark radical (a check mark with two smaller marks inside it) was accurately categorized if named an “oo”. All stimuli containing a check-mark radical were an “oo” and all stimuli failing to contain the radical were not an “oo” (satisfying the necessary condition). Additionally, any stimuli containing all the characteristics of an “oo” must be an “oo” (meeting the jointly sufficient criterion).

Hull’s (1920) application of the classical view was one of the first experimental studies of concept learning that relied on behavioral data. Indeed his study laid the foundation for further research (see e.g., Murphy, 2002) in concept learning. The interest in the classical view of concepts following Hull’s research is not surprising given the compelling characteristics that can be attributed to this view. The classical view conforms well with the way people typically communicate concepts to one another. For example, when describing a car as a vehicle a person might list features such as have wheels, doors, and an engine that propels it over roads. This description resembles the necessary and sufficient conditions required for the claim of *car*. The classical view is also parsimonious (a characteristic desired among empiricists): a concept and its natural counterpart are defined by a set of necessary and sufficient conditions. Empiricists have little difficulty understanding how a concept could be represented as a definition or as a set of necessary and sufficient conditions.

1.1.2 Prototype view

A shift in paradigm can be attributed to both inherent limitations of the classical view and several developments in the emerging field of cognitive psychology. First, the classical view tended to ignore real world concept formation, its emphasis on definitional category structures failed to capture the complexity of how human perception interacts with natural kinds to influence category/concept formation. Second, a growing body of evidence emerged suggesting that concepts were not defined as necessary and jointly sufficient features, but by typicality of characteristic features. In this view, category borders are not sharply defined, and an object belongs to a category to the extent that it resembles the prototype or exemplar of a category (Rosch & Mervis, 1975; Rosch, 1978; Brooks, 1978; Medin & Shaffer, 1978). The views that followed are often referred to as the "similarity based views" and are characterized by the notion that concepts contain and are on some level defined by shared attributes. The two similarity-based views addressed in this paper are the prototype and exemplar theories.

According to prototype theory concepts are a distribution of properties, some of which are more central or typical than others. In its weakest form, a prototype is a collection of properties composed over the typical instances of the concept. For example, the prototype for the concept *fish* might include the features swims, gills, and fins. In contrast a strong view of prototypes, holds that a prototype is a single ideal that represents a whole category. The ideal *fish* would represent all *fish*, striped, spotted, and silver, small and large, trout, bass, and angel. The prototype view shares some qualities with the classical view (e.g., see Smith & Medin, 1981). For both theories, a concept is a set of features and conditions to be satisfied by the exemplars of the concept. However, the difference between the two theories centers on how features are selected. According to the prototype theory, a person will categorize an object under the category *dog* to the extent that the object shares the same features with that category. Knowledge of an entity that barks, has fur, and wags its tail almost certainly leads to the conclusion of *dog*, whereas knowledge of an entity that simply barks leads only to the probability of a *dog* (e.g., walrus also bark). Prototypes then, are comprised of features that are probabilistically characteristic of the category but are not necessarily true for every instance. Many versions of the prototype theory also include some kind of weight that is related to each feature (Smith, Osherson, Rips, & Keane, 1988; Hampton, 1995). These weights represent how characteristic a property is of a given concept. For instance, the property of barking for the concept *dog* might be given greater weight than the property of fur.

The pioneering research of Posner and Keel (1968, 1970) provides empirical support for a strong view of prototype theory. Posner and Keele used a prototype distortion task (see also e.g., Fried & Holyoak, 1984; Smith & Minda, 1998) to create exemplars that varied around single category prototypes. Random dot patterns served as the category prototypes from which category exemplars were derived. Each category exemplar was generated by systematically varying the distance of each exemplar from its category prototype (resulting in exemplars that were either close or distant in physical similarity to the prototype). Using a category verification task, participants first learned to classify the exemplars created from the prototype dots. During this training phase corrective feedback was also provided. Immediately following the training phase some participants performed a transfer test in which they classified not only the old dot patterns (training items) but also new sets of items that were either low or high in distortion plus the never before seen prototypes; other participants performed the same test but after a one week delay. Participants received no feedback during the transfer phase of the experiment.

Posner and Keele's (1970) experiments produced several findings. First, a standard typicality effect was found; correct identification of new exemplars decreased as distortion from the prototype category increased. However, the finding of most interest resulted from the differential effect of delay and immediate testing on classification of old dot patterns and prototype patterns. Specifically, accuracy performance for old dot patterns decreased over time to a greater extent than for prototype dot patterns. For the test immediately following training, participants classified old patterns more accurately than prototypes, but for the delayed test, correct classification for prototype was reliably better than for old dot patterns. These findings suggest that following immediate classification participants have representations for both study examples and abstracted prototypes, but that memory for exemplars fades more quickly than memory for prototypes.

A strong view of prototypical representation stresses (see e.g., Posner & Keele, 1968; 1970) that category level information is stored as a single abstraction; your category for bird is represented as the single best bird, which is characterized by all the features normally found in birds. Family resemblance theorists (see e.g., Rosch & Mervis, 1975) suggest a somewhat weaker view of prototypical representation. A critical component of this view is that concepts are summary representations. A concept is represented as sets of features typically found in its category members, but some features are more relevant than others, and therefore are weighted more heavily than other features. For instance, that birds have wings is weighted heavily for the category bird, as having wings is highly typical for this category, whereas singing would carry less weight as fewer birds sing than have wings. Unlike a strong view of proto typicality, by attributing weights, the family resemblance view allows for inclusion of inconsistent features (see e.g., Murphy, 2002). It is questionable that any single prototype could represent an entire category as categories are often characterized by consistent and inconsistent attributes (e.g., birds come in all sizes; some sing while others do not). According to the family resemblance view then, concept representations are really lists of attributes, some consistent and some inconsistent. In support of this suggestion, Rosch and Mervis (1975) found that people rate some instances of a category as more representative than others and that these ratings were related to the number of typical features contained in the instance.

Categorization of new items according to a family resemblance view follows a simple additive principle. The similarity of the new item is calculated in relation to the attribute list. When the new item has an attribute that is common to the representation the item receives recognition for the attribute’s weight. If the new item fails to have an attribute contained in the representation or contains an attribute that the representation fails to have, then the new item loses the weight of that attribute. Following examination of all the items’ attributes, the weights of the items’ positive attributes are summated while negative attributes are subtracted; the new item is then categorized accordingly. If the value obtained meets the category criterion the item is judged a member of the category. If the value obtained fails to meet the category criterion the item falls outside the category.

Like the prototype view, concepts for the family resemblance view are organized according to distribution of properties, some more central or typical than others. Furthermore, family resemblance structures like prototypes are comprised of features that generally characterize an instance but need not characterize every instance. The distinction between these views is that the prototype view equates category membership for an instance with an ideal whereas for the family resemblance view category membership is determined by an instance's shared similarity to a summary representation. A characteristic of both views is that the classification of category members is dependent on similarity. To decide whether some animal is a *dog* or a *cat*, a person would compare that animal to either their prototype or summary representation and assign it to the category with which it shares greatest similarity. The prototype/family resemblance views have the advantage over earlier views of classification (e.g., see classical view, Smith & Medin, 1981) as they clearly explain typicality effects. Typical instances are those that have many traits in common with their prototype or representative member and have few traits in common with different categories.

1.1.3 Exemplar View

An alternative to the prototype view of concept representation is the exemplar view. Exemplar theorists (Brooks, 1978; Medin & Schaffer, 1978; Nosofsky, 1986) claim that instead of summary representations concepts are represented as remembered instances. Therefore a person's concept for *bird* is not an abstraction of attributes varying to a greater or lesser degree in typicality. Instead a person's representation(s) for *bird* is the set of all *bird* instances they have encountered. Like the family resemblance/prototype view categorization for the exemplar view is based on comparative similarity. However, instead of comparing a test instance to a summary representation, the test instance is compared to remembered instances of the category and any category that shares similarities. In addition to having stored numerous instances of birds a person has also stored instances from different categories that share similar attributes (e.g., *bats* also fly). When categorizing a potential bird instance all stored instances for birds plus any instance resembling a bird are retrieved. Following retrieval of the comparison categories, the category with retrieved instances that are most similar to the test instance is selected.

Note that if all the instances retrieved belong to the same category then categorization of the test instance is a straightforward automatic process. The attributes of the test instance are compared those of other retrieved category members and if the required number of attributes are shared then that test instance belongs to the category. However, if the retrieved instances belong to different categories, then the categorization of the test instance becomes more complex. The similarity of the test instance to each retrieved member must be computed. The similarity scores are then combined over members of the categories, and finally the category having the highest similarity score is chosen.

Exemplar theorists (Medin & Schaffer, 1978; Nosofsky, 1986) propose that memory for old instances alone can account for the transfer patterns found in Posner and Keele's (1970) research. According to Medin and Schaffer (1978), the advantage for old items on the immediate testing results from the identical match of these items to instances stored in memory. However, with delay, memory for specific instances has degraded, so that identification of an old item is more problematic. For prototype items, outcomes are reversed. At immediate testing, prototypes are not an identical match to memory for old items so categorization is more difficult. With delay, memory for prototypes is stronger because they have similarity to many items and specific information for old items has been lost. As specific information is lost, critical cues to the correct identity of the item are also lost therefore the weight of recollection will be for the item that appeals to a general memory, which is the prototype item.

Exemplar models appear to be more sensitive to within category correlations than prototype models (e.g., see Malt & Smith, 1984; Medin, Altom, Edelson, & Freko, 1982). For instance, exemplar models can easily represent the idea that small birds are more likely to fly than extremely large birds. This is because each instance is individually represented in memory and can be retrieved as stored. However, a clear problem for prototypical views is an inability to account for continuous dimensions. This is because new features are simply added to existing features lists. This leads to the rather untenable situation in which a bird may be concurrently both small and large. Over a series of experiments Medin, et al. (1982), showed that exemplar-based models are sensitive to within category feature correlations. In one experiment, they tested sensitivity to within category correlations by having participants study items from a category containing disease features. Two dimensions in the category were always highly correlated (e.g., splotches on the skin were always highly correlated with high red blood cell). At test participants consistently judged correlated items higher than items that failed to preserve this relationship (e.g., splotches on skin correlated with white blood cell count). These results were further supported using more than one category and artificial items (Medin et al. 1982), and even natural categories and items (Malt & Smith, 1984). Based on this research evidence, it would appear both theoretically and empirically that at least for within category correlations exemplar based models hold the advantage.

1.1.4 Knowledge View

A recent criticism of prototype and exemplar views involves their one-dimensional focus on properties and overemphasizing similarity as a means of explaining why we have some categories and not others (see e.g., Murphy & Medin, 1985; Storms & De Boeck, 1997). Research demonstrates that people have extensive knowledge for familiar concepts that property lists fail to capture (see e.g., Murphy, 2002). Asking people to list relevant properties for a concept produces functions, beliefs, relations among objects, subordinates, superordinates etc. In other words the properties listed for concepts extend far beyond the simple property listings afforded by perception alone. Regarding similarity, these approaches fail to capture why some similarities matter while a large number of others do not (see e.g., Murphy & Medin, 1985). Similarity is too flexible to explain conceptual uniformity: Any two objects can be arbitrarily similar or dissimilar by changing the criterion for what counts as a relevant property. Thus, similarity is only useful to the extent that principles determining what counts as a relevant property are specified. Such principles are believed to arise from background knowledge that people have about the world (Murphy & Medin, 1985).

In response to apparent insufficiencies to property listings and unconstrained similarity matching, the theory-theory or knowledge view was proposed (see e.g., Murphy & Medin, 1985; Murphy, 2002). According to this view concepts are mental representations that serve as building blocks (Medin & Ortony, 1989) for human thought and behavior. Concepts may not necessarily have a real world instantiation (e.g., chimeras), and people may construct rather than discover structure on the world (see e.g., goal derived categories, Barsalou, 1983,1985).

The knowledge view marked a shift from a perceptual emphasis in categorization toward a more theoretical/inferential account. Within this approach, no single characteristic defines concepts or categories; there are in fact many kinds of concepts and categories, including fuzzy, natural, abstract and artificial ones (Medin, Lynch & Solomon, 2000). Solomon, Lynch, and Medin (1999) for example, describe concepts as mental constructs that serve multiple functions such as categorization, learning, reasoning, and communication. Barsalou (1987) argues that concepts are unstable, likely to change between and within participants over time depending on context and prior experience. Some researchers define categorization within the context of purpose and declare that categorization is primarily for inferring *unseen* features (Kruschke, 2005) or making accurate predictive inferences (Anderson, 1991). As the brief review of the knowledge literature that follows suggests, there are many ways in which prior experiences can influence categorization. The remainder of this section will explore these influences in greater detail as well as examine recent criticism of the approach.

In their seminal paper, Murphy and Medin (1985) contrasted the prototype and exemplar views with a theory-based approach. Murphy and Medin argued that while intuitively appealing, similarity-based approaches were incomplete. Specifically similarity-based approaches do not provide enough constraints. The categorization of even the simplest of objects is influenced by more than simple feature comparisons (see e.g., Heit, 1994; Murphy, 2002). Handbags and wheelbarrows, for instance share many commonalities: they carry things weigh less than 2000 pounds, and are inanimate. If similarity alone accounts for categorization, why are handbags and wheelbarrows not categorized as like objects? Furthermore, the learner is an active agent in the categorization process. Handbags do not come labeled with tags such as *opens* and *carries object*. It is only through interaction with the object that the learner infers and understands the object that is being categorized (e.g., see also Ross, 1997). In other words, concepts are not rigid structures but are changing as new information is incorporated. Finally, people have knowledge of causal relations between features not contained in a list of features. For example, its reasonable to assume that larger birds are more likely to live in tree tops and smaller birds on the ground as larger birds are more likely to resist the severe elements that living in tree tops entails.

According to Murphy and Medin (1985) similarity-based approaches are unable to address these critical influences of knowledge on concept formation and categorization. In contrast, the theory-theory view (Gopnik & Wellman, 1994 used this term in order to describe Murphy and Medin's approach) posits that the organization of concepts is knowledge based and theory driven, while categorization is an inference process, not a similarity judgment. Early evidence for the theory-based approach can be found in Barsalou's (1983) research into *ad hoc* categories. Ad hoc categories are created on the fly and cannot be interpreted as fixed structures. For instance, *things to take out of a burning house*, is a category that does not conform to any specific, pre-existing category and, in a very real sense, is a different category to different people. Barsalou found that while *ad hoc categories* conform to a graded structure and show typicality effects, these categories do not show a family resemblance structure; rather the categories are a collection of apparently dissimilar members (e.g., children, paintings, jewelry).

Barsalou describes these categories as goal oriented. This research and additional research by Barsalou (see also 1982; 1985; 1991) not only demonstrated that categories can be generated as needed, but that classification is based on more than the simple matching of an object's properties to those of a stored representation(s). People use their prior knowledge of categories to infer and make new causal relations. Indeed, matching members to their respective categories appears to require the right explanatory relationship and prior knowledge appears to be the mechanism by which those connections are made.

Some researchers (Ahn, Brewer, & Mooney, 1989, 1992; Medin, Wattenmaker, & Hampson, 1987; Wattenmaker, Dewey, Murphy, & Medin, 1986; Wisniewski & Medin, 1994b) have examined how people use their prior knowledge of categories to represent and interpret what they observe. For instance, Wisniewski and Medin (1994b) demonstrated how categorization is sensitive to prior knowledge and how that knowledge influences understanding of the category members. In their studies, participants learned the same sets of children's drawings but with different category labels. Each group learned two sets of drawings: For participants in one group the drawings were labeled *drawn by city children* or *drawn by rural children* while for participants in another group the drawings were *labeled drawn by gifted children* or *drawn by normal children*. Wisniewski and Medin found that how participants interpreted the drawings was highly influenced by the category labels they had been exposed to. For example, participants exposed to the so-called gifted drawings were far more likely to describe the picture as having an *unusual* and *creative* quality than participants exposed to the same picture with a category label of normal. Clearly, prior knowledge can have a dramatic affect on interpretation of objects.

Prior knowledge may also play an important role in how critical features are selected. Specifically, during category learning prior knowledge may lead one to selectively attend to those features that are particularly relevant to the categorization of the object (Murphy, 2002; Murphy & Medin, 1985; Murphy & Wisniewski, 1989).

Pazzani (1991) investigated the role of selective attention by having participants learn categories of balloons using pictures of adults or children performing actions on deflated balloons. The pictures varied along several dimensions including: who was performing the action (child or adult), the type of action (whether the balloon was being stretched or dipped in water), size (small or large), and color (purple or yellow). Participants were instructed to learn one of two categories of balloons, labeled *one that inflates* or simply *Alpha*. In any one condition the categories were described according to a simple disjunctive rule such as the *balloon must be stretched* or *inflated by an adult*. In other condition participants were exposed to a conjunctive rule the *color must be yellow and the balloon must be small*. Past research had demonstrated that in comparison to the disjunctive rule people found it easier to learn conjunctive rule. Pazzani reasoned that the conjunctive advantage might be reversed if the disjunctive rule cued participants' past experiences with inflating balloons. That is the disjunctive rule would activate prior knowledge consistent with inflating balloons (e.g., stretching balloons results in easier inflation and adults are better able to inflate the balloons than children) making categorization easier. Under these conditions Pazzani found that category learning for the *inflated* disjunctive category group was much faster than for the category labeled *Alpha* conjunctive group. These findings suggest that participants in the inflated condition used prior knowledge associated with the rule to selectively attend to those features most relevant to categorization of the object.

Prior knowledge can also function to assist or facilitate category learning and concept formation (Medin & Schwanenflugel, 1981; Murphy & Allopenna, 1994; Wattenmaker, Dewey, Murphy, & Medin, 1986). In this view, learning about certain kinds of category structures is influenced by the type of information and expectation one has in relation to the structure. Medin and Schwanenflugel (1981) differentiate two types of category structures, *linearly separable* and *nonlinearly separable*. This research was important for testing differences between prototype and exemplar theories. Linearly separable categories, which are most clearly associated with prototype theories, allow for independent summation of category features. Criterion for category membership is met if the majority of features for the category candidate match the target category. For example, if a category's candidate barks, wags it tail, and has four legs it will likely fall into the category of *dog*, as these features are all typical of dogs. In contrast, nonlinear category structures, which are mostly associated with exemplar theories, cannot be learned by simple summation of category features. With this type of structure relying on individual features alone in order to determine category membership is less helpful; people must form specific groupings or relations for features and categories. Results from previous research (see e.g., Rosch & Mervis, 1975) suggested that linearly separable categories were easier to learn than nonlinearly separable categories. However, Medin and Schwanenflugel found that when participants were given instructions that promoted learning the (nonlinear) type of structure they were exposed to no differences were found in participants' ability to learn linear and nonlinear category structures. Focusing on information relevant to the type of category structure facilitated the category learning.

Wattenmaker, Dewey, T.Murphy, and Medin (1986) extended Medin and Schwanenflugel's (1981) research by examining how prior knowledge influences the learning of linear and non-linear separable categories. In their study, participants were placed in one of two groups, the trait or control. In the trait group, labels for stimulus dimensions cued participants' prior knowledge of specific personality traits. For example, some labels cued participants' prior knowledge for behaviors that were either honest (e.g., returned lost wallet) or dishonest (e.g., pretending to enjoy shopping). In the control group the labels for stimulus dimensions cued four unrelated traits (e.g., talkativeness, cooperativeness, cautiousness, honesty). Under these circumstances, there were no coherent cues that would promote retrieval of specific personality traits. The main finding was that coherent cues for personality traits facilitated learning of categories; participants in the trait group learned the categories reliably faster than participants in the control group. Thus, making the task more meaningful assisted category learning. Furthermore, for participants in the trait group, linear separable categories were easier to learn, suggesting participants learning these categories used prior knowledge to distinguish honest and dishonest behaviors. In contrast, participants in the control condition found nonlinear categories easier to learn. Having no prior knowledge, they likely formed a specific configuration of traits (e.g., honesty and cooperativeness might be associated with the category, but cooperativeness alone as not). Overall, the results suggest that learning was most influential when the structure to be learned matched the structure associated with prior knowledge.

Finally, some researchers (Heit, Briggs, & Bott, 2004) report limitations and boundaries around facilitation effect of prior knowledge. Heit, Briggs, and Bott (2004) conducted three experiments addressing how observations and multiple sources of prior knowledge interact in category learning. In their Experiments one and two, learning was faster for key features, which were predictable on the basis of prior knowledge, than for irrelevant features. Moreover, this advantage increased as more observations were made. In their Experiment three, however, presenting feature information that went against prior knowledge led to little overall effectual use of prior knowledge. Thus, when information is inconsistent with prior expectations, the usual facilitation effects associated with prior knowledge may not occur.

1.1.5 Criticism of the knowledge based view

Criticism of the knowledge-based view often involves attacking basic tenets of the approach. Rosenblit and Keil (2002), for instance (see also Komatsu, 1992) argue that, like similarity-based approaches, the knowledge view does not offer enough conceptual constraints. As discussed earlier the central tenet of the view (Murphy & Medin, 1985) is that concept formation and the categorization process are driven by people's theories of how properties are related. Keil (2003) notes that the type of theorizing the average person performs are in stark contrast to scientific theorizing and as a consequence are extremely difficult to quantify. Scientific theorizing involves examining causal relationships by deconstructing the principles of these relationships into functional units and then analyzing how these units interact to produce the event in question. Once a causal relationship is understood, further exploration may be unnecessary. However, theorizing for the layperson is "potentially unbounded". A person's theory about how birds fly, for instance, might include “feathers help birds fly”. However, this explanation in turn forces one to ask why feathers are necessary for flight. Why not scales? In other words, unlike scientific theorizing, which has known methodological constraints; the average person is not bound by these constraints. As key principles among relationships are rarely thought out, causal relationships for the average person are never complete; this may limit the ability to decompose an event into its functional units for further analysis. Additionally, there are likely large individual differences in how people understand causal relationships.

Research by Rozenblit and Keil (1997) rules out the average person using reasoning strategies characteristic of scientific theorizing as a method of imposing these constraints. Over several studies, they demonstrate that people confidently believe they know how things work, but when challenged are forced to concede that their understanding is superficial and even illogical. In their experiments, participants were presented with an item, and asked to rate their knowledge of the item’s mechanical functioning on a seven-point rating. Following the knowledge ratings, participants were asked to give a detailed description of how each device works and why the device functions the way it does. Next they were asked a deeper question, one that probed their understanding of the objects fine mechanics (also followed by a rating judgment). Finally, they were given an expert explanation as to how the object works and then re-rated their understanding of the object. Rozenblit and Keil report that participants in these experiments show a large reduction in their knowledge ratings and furthermore are overwhelmingly astonished by their initial overestimation of their understanding.

Rosenblit and Keil (1997) suggest that conceptual theorizing is on the whole somewhat shallow. Unless some kind of external structure is imposed, as is the case with science for instance, people's use of knowledge for making causal connections may be far coarser than some theories might suggest. However, unlike the type of theorizing that occurs in science where the mechanics and role of the concept are clearly defined, the average person has no such constraints. Lutz and Keil (2002) further suggest that what remains of the knowledge view, is that people have knowledge of a higher but more shallow level of causal information. For instance, at a relatively early age children learn what properties of a domain are important for classification. They understand that color is important for distinguishing kinds of plants, but not for distinguishing kinds of cars. Children also develop a clear understanding of the importance that domains of expertise play in further understanding the world around them. Teachers, parents, and friends all vary in domains of expertise, and knowing which of the domains to access demonstrates a high level, though shallow, understanding of the domains themselves.

1.1.6 Summary

It is clear that categorization theories explain many basic aspects regarding influences of knowledge on concept learning. However, some questions still remain. Empirically based models, for instance, often use artificial stimuli, which are removed from any knowledge, and yet people classify reasonably well under these conditions. Clearly, the knowledge approach has not fully replaced similarity-based views, as these views give reasonably good explanations as to how statistical information is learned.

A clear understanding of how knowledge changes the learning process has not yet been established (see e.g., Murphy, 2002, for review). For example, some explanations of category learning suggested (Murphy & Allopenna, 1994) that people did not focus mental resources on learning statistical properties but instead once they identified the category, knowledge of the category became the category representation. For instance, once you have determined an object was a *cat*, you didn't then devote your attention to learning the properties of a *cat*, but simply represented the category as a *cat*. This assumption was later re-evaluated when some researchers (Spalding & Murphy, 1999), demonstrated that knowledge might facilitate learning of statistical properties.

Knowledge theorists (Murphy, 2002) have largely focused on how prior knowledge greatly speeds category learning or does not impair statistical learning. Only recently have theorists begun an investigation into how information incongruent with prior experiences may interfere with the retrieval of category information. Heit et al. (2004) showed that feature information that went against prior knowledge led to little overall effectual use of prior knowledge. In that study, knowledge pre-empted learning of empirical properties. The present study extends prior research by exploring how abstract item information interacts with deeply embedded prior knowledge structures when that knowledge varies in specificity. Specifically, one focus of the present study is on how prior knowledge interacts with learning of abstract information over taxonomic structures. Before addressing the present study a description of taxonomic structures is presented as well as an overview of taxonomic research central to the category literature.

1.2 Taxonomic structures

Many adult concepts can be represented in taxonomies – hierarchical systems in which concepts are differentiated into varying levels of abstraction (e.g., *musical instrument, wind instrument, flute*) related by class inclusion (a *flute* is a *wind instrument*, and a *wind instrument* is a *musical instrument*). Indeed, most natural kinds (e.g., *whale*, *tree*) and artifacts (e.g., *flute*, *truck*) are generally believed to fall within taxonomies. Moreover, in real world contexts, concepts are rarely learned as explicitly contrasting sets existing completely outside of known taxonomies (i.e., one might not learn *cats* vs. *dogs* without also learning that both are types of animals, and that both include more specific subcategories).

Typically, conceptual taxonomic structures consist of both vertical and horizontal relational links. Vertically, concepts are taxonomically related when they are hierarchically organized from the more to less inclusive levels or from less inclusive to the more inclusive ones (e.g., *mammal* to *cat*, or *Persian* to *cat*). Horizontally, taxonomic structures relate a concept of one hierarchical level to another concept at the same level (e.g., *cat* to *dog*).

It is generally assumed that taxonomic architecture functions both efficiently and economically. Properties shared by concepts at the higher and more inclusive level are transferred to the concepts at the lower level but not necessarily in the reverse. For example, properties true to *animal* (a superordinate level concept) such as breathing are also true of *cat* while properties true of *cat* or types of *cat* such as purring, are shared by all *cats* but not necessarily by all other animals. Thus, hierarchical structures store properties of concepts in an economical way. Furthermore, as suggested by the example, taxonomic structures promote efficiency of learning. Knowing that all animals have skin allows you to infer all *cats* have skin even though you may never have been exposed to one. Similarly knowing that all *cats* purr naturally leads to the conclusion that all *lions* purr.

Though potential advantages of taxonomic hierarchies are apparent their status is not entirely clear. Several questions arise with respect to taxonomic hierarchies including: do they have psychological reality? Is there a privileged level, which is optimal for storage and communication of information? The remainder of this section will explore these questions.

1.3 Conceptual Hierarchies.

The use of hierarchical structures is evident across a large number of domains including anthropology, biology, and psychology. Within the domain of psychology the use of hierarchical structures as a means of explaining various phenomena is particularly evident in the field of cognition. Exploring the organization of hierarchical representation, for example, began to dominate cognitive psychology during the 1960's. However antecedents can be identified as early as the 1920's. Lurias's model of brain organization (published in 1970, but crafted in 1920's and 30's) had three functional units: one for programming and self-regulating activity, one for processing and storage of information, and one for regulating consciousness. With the emergence of cognitive psychology in the 1960's and the subsequent trend toward focusing on the nature of mental representation, cognitive models at this time (e.g., Collins & Quillian, 1969; Rosch & Mervis, 1975, 1976) often included a hierarchical structure.

The rise of hierarchical models in cognition prompts an important question: Do hierarchies have a real existence in terms of conceptual reality or do they simply reflect conventions adopted by researchers? This question is not new and has been debated in other domains. In biology, for example, where taxonomies are often defined within the restrictions of set laws, Linnaeus (see e.g., Denton, 1985) took a decisively realist approach to the existence of biological species. With the advent of evolutionary theory and the belief in variability of species a view arose which suggested that concepts of species exist not as real universals, but as human ability to impose systems of classification. As such, classification systems were seen as a convention and by extension so were hierarchical structures (Denton, 1985).

1.3.1 Psychological status of taxonomies

Early views of taxonomic structures posit (e.g., Collins & Quillian, 1969) that they are relatively stable knowledge structures characterized by property inheritance. One of the earliest models of taxonomic formation was inspired by the notion that computers could develop human characteristics. Quillian (1967) introduced a computer program based on a hierarchical network of semantic memory in which concepts were represented as interconnected nodes. The network is hierarchical in that higher-level superordinate concept nodes have connections to the lower level basic and subordinate conceptual nodes. For example, the concept *animal* is connected to the lower basic level concept node of *dog*, which in turn is connected to the even lower subordinate level concept nodes of *pit-bull* and *German shepherd*. The network itself follows a principle of cognitive economy. Properties true of all animals, like reproduction and breathing are stored at the *animal* node. Similarly, properties generally true of an entity (e.g., barking) are stored with the particular concept they represent (e.g., *dog*). A property does not have to be true of all lower level concepts to be stored with a higher-level concept. Fur, for instance, is stored with the concept node of *dog*; those instances of dogs that do not have fur would have their properties stored at their individual concept nodes. Category membership is determined by the concept's position in the hierarchical network. The node *dog* does not store the information that dogs are animals; instead membership is determined by first activating the concept *dog*, followed by the activation of *mammal* and finally *animal*.

The suggestion that properties are stored efficiently at their concept node and that it takes time to move over the network produces a number of testable predictions. For example, when traveling over concept nodes in the hierarchy, the time needed to verify concept features should increase as the distance between one concept and another concept increases. Therefore, people should be faster to confirm that all *dogs* bark than to confirm all dogs have fur and faster to verify that all *dogs* have fur than all *dogs* have skin (stored at the animal node). Collins and Quillian (1969) later found support for these types of predictions.

Rips, Shoben, and Smith's (1973) research challenged Collins and Quillian's (1969) notion of taxonomies as pre-stored memorial structures. Rips et al. explored the influence of typicality ratings on class inclusion judgments. In their study, participants completed a timed sentence verification task in which they judged whether category members were typical or representative of their superordinates. For example, they might judge how typical the category *robin* is of the category *bird* vs. how typical the category *peacock* is of the category *bird*. Rips et al. also examined reaction times for typicality ratings over multiple superordinate levels (e.g., how typical a *robin* is of *animal*, vs. how typical a *robin* is of *avian*). A memorial-based view of taxonomic hierarchies holds that taxonomic relations have a veridical representation in memory. Because the categorization process is a strict sequential route, there should be no difference in categorizing two similar categories within their respective domain (e.g., *robin* or *peacock* as *bird*) and leaping over a superordinate (e.g., *cat* as an *animal*) should result in slower reaction times than categorizing the same object in relation to its closest superordinate (e.g., *cat* as a *mammal*). Yet Rips et al. found that response times for same category members varied, in particular those categories that could be interpreted as more typical or representative (e.g., *robin*) of their superordinate category (e.g., *bird*) were verified faster than categories viewed as less typical (e.g., *peacock*). Additionally, verifying the less typical sentence, *a cat is a mammal* was slower than verifying the more typical sentence a *cat is an animal*. Rips et al. concluded that a strict memorial representation of taxonomic structure was unsupported; participants were more likely to make judgments based on how typical or representative the category was of its superordinate (see also Murphy, 2002 for an overview of this topic).

Research by Sloman (1998) challenges the notion that people always invoke class inclusion relations. Over a series of experiments, participants reasoned and evaluated the strength of various inductive arguments involving natural, social, and artifact kinds. Below is an example of type of argument structure.

(A) <u>All bodies of water have a high number of seiches.</u>
All lakes have a high number of seiches.

(B) <u>All bodies of water have a high number of seiches.</u>
All reservoirs have a high number of seiches.

For each of Sloman's arguments, the conclusion category (e.g., lakes) is incorporated into the premise category (e.g., bodies of water). The arguments stress that a property true of all members of the premise category is also true of all members of the conclusion category. For each argument, participants gave probability ratings that assessed their perception of how true the conclusion category was of the premise category. The findings were conditional on participants' previously agreeing that the conclusion category was included in the premise category. In this experiment, not one participant rated all of the arguments as definitely true, and no one argument was rated as definitely true by all participants. However, the similarity between the premise and the conclusion category did affect probability judgments. Participants rated argument "A" as more probable than argument "B", which suggests that participants abandoned class inclusion relations when reasoning about these categories in favor of similarity relations.

Unlike previous research (see e.g., Hampton, 1982) examining class inclusion relationships, the argument structure for many of the items in Sloman's (1998) experiments were unambiguous. This lends strength to Sloman's (1998) argument that the experimental findings from his studies support a similarity based reasoning strategy. Moreover, in separate but related experiments, Sloman used the same argument structure but made the class inclusion relationship more explicit and thus more accessible. Below is an example of a strengthened argument structure.

(C) All lakes are bodies of water

<u>All bodies of water have a high number of seiches.</u>

All lakes have a high number of seiches.

Under these conditions participants rated conclusions as certain, suggesting that when the relation is made explicit it is also made more accessible. Sloman concluded that even though people are capable of correctly using inclusion relations they do not always explicitly represent them.

While Sloman's (1993, 1998), model contradicts simple views of taxonomies as permanent knowledge structures, the model itself is nevertheless driven by the conjecture that class inclusion inferences are based largely on taxonomic connections between concepts. Arrays of features representing concepts are compared to each other, and inferences are strong to the degree that premise and conclusion concepts are made salient by shared attributes. However, recent research (Medin, Lynch, Coley, & Atran, 1997) challenges both the importance of similarity when making inductive inferences, and the notion of hierarchies as single structures characterized by a sequence of nested categories (e.g., a specific type of *bird*, *robin* belongs to the categories of *bird*, and *animal*). Medin and colleagues (1997) examined categorization among different types of tree experts. They found that landscapists used one hierarchical structure for categorization but a different hierarchical structure for reasoning. Specifically, landscapists sorted trees into categories on the basis of goal-relevant and practical reasons. Medin et al. suggested that years of experience resulted in goal derived categories becoming embedded in memory. However, when reasoning about whether a natural attribute true of one tree was true of another, inferences made by landscapists followed a pattern more in accordance with taxonomic relations. In contrast to landscapists, park maintenance workers based their sorting almost entirely on ecological factors. That is, their decisions were driven almost entirely by practical issues (e.g., what trees best fit a given region). Medin et al. have shown not only that people use different types of organizational structures within a given domain, but also that background knowledge strongly influences what structures will be used within the domain.

In a series of experiments, Ross and Murphy (1999) used the domain of food to demonstrate that people classify foods into both taxonomic and non-taxonomic categories. They found that some types of foods (e.g., beef) were classified taxonomically as a type of meat and non-taxonomically as a main course. Furthermore non-taxonomic categories often included categories from different domains and thus failed to follow a strict category induction pattern. For instance, dairy products (e.g., milk), fruits (e.g., banana) and meats (e.g., fish) were classified under the category of healthy foods, even though each of the foods formed independent taxonomic structures. One conclusion drawn from Ross and Murphy's research is that people form non-taxonomic categories because their reasoning skills extend beyond simple similarity judgments (see also e.g., Murphy, 2002) to include complex reasoning skills that are goal driven (e.g., know that people eat healthy foods to avoid illness). Ross and Murphy also concluded that non-taxonomic categories are established in memory and are important to the category inference process.

In sum, research suggests that when taxonomic connections are made, it is the salient taxonomic inferences that are more likely to be extracted (Sloman, 1993, 1998). Additionally, knowledge heavily influences the use of organizational structures, with experience guiding the use of taxonomic and non-taxonomic inference (Murphy & Ross, 1999). The next section examines and describes the core research question cognitive psychology has predominately explored with respect to taxonomies, namely the question of a privileged taxonomic level.

1.3.2 Basic level superiority

The notion that the basic level of the taxonomic hierarchy has advantageous characteristics not true of other taxonomic levels is largely driven by the research of Rosch and colleagues (1976a). In their early experiments, Rosch et al. used free naming/category-verification tasks to demonstrate that basic (intermediate) levels of the taxonomic hierarchy have a privileged status relative to subordinate (lowest level) or superordinate (highest level) categories. In the naming task, participants were presented with a series of pictures in rapid succession and were asked to write down the name of the object depicted in each picture. There were two clear findings; first, participants used basic-level names (e.g., *chair*, *hammer*) more frequently to identify an object than subordinate (e.g., *kitchen chair, ball peen hammer*) or superordinate (e.g., *furniture, tool*) level names; second, of the three levels superordinate names were used least frequently. In the category-verification task, participants heard a category label representing superordinate, basic or subordinate level categories and then indicated whether a picture shown after a brief delay was an instance of the category. Results showed that objects were often verified faster at the basic level than at the subordinate or superordinate levels. Based on these findings, Rosch et al. argued that people access the basic level first and then access the subordinate and superordinate categories. Rosch et al. also suggested that the basic level advantage arose because members at this level have attributes that are both distinctive and informative whereas members of other taxonomic levels have only one or the other of these attributes. That is, category members at the basic level have distinctive features which help differentiate them from members of other categories, but these same features also overlap with members of their own category, thus helping to define their category membership (e.g., all *birds* have *wings* but only a few members of other categories have *wings*). In contrast, superordinate categories though distinctive are not informative (e.g., *furniture* is very different from *vehicle* but what are the features that define *furniture*?) and subordinate categories, though informative are not distinctive (e.g., there are many features which define *dining room table* but few to distinguish it from a *kitchen table*). Collectively these assumptions became known as the differentiation hypothesis (see also Murphy & Brownell, 1985).

As a result of this and related research (e.g., Rosch & Mervis, 1975; Rosch 1978), Rosch argued that concepts follow a natural construction. Objects in the world are disposed toward forming clusters of correlated attributes; category groups and concepts naturally arise out of these clusters. For example, a bear is more likely to have fur than scales; alternatively a snake is more likely to have scales than fur. The implications of Rosch's argument are clear: in order to understand concept formation, the focus should be on natural categories, and advancement of categorization only occurs by studying natural categories.

It is also important to note that in addition to emphasizing structure, Rosch (1978) also recognized the role of the perceiver. Rosch stressed that it is the relationship between the perceiver and the world that specifies the basic level. As a consequence an individual's expertise in a domain can play an important role in specifying the nature of the basic level. Rosch et al. (1976) observed that participants who had greater knowledge in a domain answered their questions with greater specificity than participants who had limited knowledge. Based on this observation, Rosch et al. noted that the experts in their field know more than novices; the contribution of the perceiver to the categorization process then could be assessed by systematically varying levels of expertise within a given domain. Under these conditions, Rosch et al. reasoned that the more specific subordinate level categories might substitute for basic level categories as the level of first access.

In one of the few papers to examine category learning within taxonomies, Murphy and Smith (1982) used artificial materials to examine the nature of the basic level advantage. Specifically they examined whether Rosch's et al. (1976a) finding of a basic level advantage had been influenced by the following (1) word length, subordinate words were longer than basic, (2) familiarity, basic level names tended to be more common than subordinate level names, or (3) saliency of features, features at subordinate level were not as distinctive as those at the basic level. Another factor influencing Murphy and Smith's research involved resolving an ongoing debate within the category literature. As suggested earlier, Rosch et al. (1976a) claimed that the basic level first advantage arose because members at this level have both distinctive and informative attributes. This view contrasts with that of Anglin (1977) who holds that order of learning determines basic level superiority. Basic level categories are learned earlier than other levels. Therefore basic levels have the advantage of prior experience.

In order to control for factors such as word length, familiarity, saliency, and prior experience, Murphy and Smith (1982) used artificial materials. Stimuli used included pictures of highly distinctive novel tools, some of which characterized the basic level categories, and others the superordinate and subordinate level categories. Basic level categories were both informative and distinctive, superordinate categories were distinctive but uninformative, and subordinate categories were informative but not distinctive. Category names for each taxonomic level were of equal length and participants were informed of all relevant features. Order of learning was further controlled for by varying category presentation; if order of learning determined level superiority then the category level learned first would be categorized fastest. For the training phase, participants learned each category independently, that is they learned one category (e.g., basic level), were tested then learned the next category. After learning all the category levels they were given a timed categorization task, in which they heard a category name, and viewed a picture.

Using this procedure Murphy and Smith (1982) found that neither order of learning nor familiarity played a role in basic level superiority. Furthermore the basic level advantage was found regardless of which category level was learned first. Generally, basic level categories were learned fastest followed by subordinate then superordinate. In subsequent experiments using similar methods with slight modification Murphy and Smith found support for Rosch et al.'s assertion that the basic level category advantage arose as a result of their distinctive attributes. However, this claim was qualified by the need for attributes to be perceptual in character. In sum, Murphy and Smith found tentative support for the basic level advantage and differentiation hypothesis and little support for the order of learning view.

Murphy and Brownell (1985) tested specific qualities of the differentiation hypothesis and found additional support for this view. They used a picture categorization task in which the typicality of stimuli was varied. They suggested that although typical subordinate level items are not distinctive (e.g., *robins* share many similar attributes with other *birds*), atypical basic level category items, for instance *penguins*, are distinctive because they share few characteristics with other members of their category (e.g., *penguins* have many attributes in common that are also distinct from those of other *birds*). Furthermore, as penguins are members of the subordinate level category, they already are informative. As predicted by the differentiation hypothesis, compared with pictures of typical subordinate items, pictures of atypical subordinate category items were categorized faster as a member of their own category than as a member of their basic level category. For example, participants were faster to verify a *penguin* as a *penguin* than as a *bird*, but slower to categorize a *robin* as a *robin* than as a *bird*. These findings are important for several reasons. First, they suggest that the basic level is fluid. Under the right conditions, all the advantages (e.g., speed etc.) usually attributed to the basic level can also be attributed to subordinate level categories. Second, the view (see e.g., Anglin, 1977) that Rosch et al.'s (1976) previous differentiation findings can be attributed to other determinants (e.g., familiarity, frequency etc.) is further minimized. The word *robin* is clearly encountered and used more often than the word *penguin*. In subsequent experiments, Murphy and Brownell found further evidence for the differentiation hypothesis. By presenting category members in contrasting contexts (e.g., a *robin* shown in the presence of a *hammer* or a *car*), shared category members can be removed from comparison (that is, robin was no longer being compared to other similar *birds*, for instance *jays*). Murphy and Brownell demonstrated that under these conditions participants were in fact faster at identifying a subordinate member as a subordinate than identifying a subordinate member as a basic. From these findings it would appear that lack of distinctiveness is responsible for previous reaction time deficits found for subordinate level members. That is, by increasing distinctiveness, having to distinguish from similar category members is no longer a problem.

Tanaka and Taylor (1991) extended previous research into basic level superiority (e.g., Murphy and Brownell, 1985; Rosch et al. 1976) by investigating the affect of expertise on taxonomic structures in expert bird watchers and dog breeders. Using similar procedures to that of Rosch et al. (1976), Tanaka and Taylor tested each expert in both the *bird* and *dog* domains. Consequently, when tested within their own domain bird watchers and dog breeders were experts, but when tested outside their domain they were novices. Participants completed three tasks: feature listing, free naming, and speeded categorization. Consistent with previous research (e.g., Rosch et al. 1976) experts in their novice domain listed many more features at the basic than at the subordinate level. However, in their domain of expertise, participants listed many more features at the subordinate level than at the basic level. For example, *dog* experts listed as many features for the basic level category of *bird* as *bird* experts, but failed to list as many features for the subordinate level category of finch. In the free naming task, participants viewed pictures of objects and responded with the first name that came to mind. As novices, participants generated basic level names the majority of the time. However, in their expert domain production of category names was dependent on type of expertise. Specifically *bird* experts generated subordinate names almost equal to their production of basic level names as novices, whereas *dog* experts also used subordinate level names though less frequently. Both bird and dog experts outperformed novices in production of subordinate level names. Finally, in the speeded categorization task, participants viewed a category name followed by a picture. The results showed that experts in the novice domain were fastest at the basic level and slowest at the subordinate level (thus replicating Rosch et al. 1976). However, in the expert domain, categorization performance was equally fast at the basic and subordinate levels.

Similar to Murphy and Brownell's (1985) research, Tanaka and Taylor's results question whether any single taxonomic level is *basic*. Domain expertise, like categorization of atypical subordinate categories, enhanced the speed at which subordinate level objects were accessed. Indeed, domain expertise caused subordinate level categories to become as accessible as basic level categories. It is important to note Tanaka and Taylor's (1991) results extended the research of Murphy and Brownell (1985) in several ways. First, feature listings by experts demonstrated that subordinates *were* distinctive; as experts, participants listed many features for subordinate categories, but as novices they listed fewer for basic level categories. Second, in comparison to the novices, experts were not bound by basic level constraints. They were as comfortable categorizing items at the subordinate level as they were at the basic level. It is likely then that the internal structure of taxonomies is not determined solely by the correlation structure of the environment (e.g., Rosch & Mervis, 1975; Rosch 1978) but on some level also reside in the mind of the perceiver (see e.g., Lynch, Coley & Medin, 2000; and Medin, Lynch, Coley & Atran, 1997).

Interestingly, Murphy and Wisniewski (1989) demonstrated that under the right conditions superordinate-level categorization could also supersede the basic level advantage. Previous research using tasks designed to differentiate superordinate and basic level performance found that participants were slower to categorize superordinate pictures (Murphy & Brownell, 1985; Rosch et al. 1976; Smith, Balzano, & Walker, 1978); slower to write superordinate names (Smith, Balzano, & Walker, 1978); and slower to categorize artificial superordinate objects (Murphy & Smith, 1982). Common to all this research was the presentation of objects in isolation. Murphy and Wisniewski (1989) suggested that superordinate categories might contain "relational information" that connects objects in a given category. As a result, presenting objects in isolation may mask a superordinate level advantage. In order to test their hypothesis, Murphy and Wisniewski presented objects either individually or in groups. They found that presenting objects individually resulted in the typical basic level advantage (e.g., a *hammer* was categorized faster as a *hammer* than as a *tool*). However, when objects were presented in a contextual scene (e.g., *hammer* presented along with other *tools*, a *workbench*, etc.), participants categorized the items as fast as other superordinates at the basic level items. In sum, evidence does not always support a basic level (Rosch et al. 1976) advantage; people do not necessarily access the basic level prior to accessing the subordinate or superordinate levels.

1.3.3 Summary

Research on taxonomy has largely focused on two issues, the psychological status of taxonomies, and the level at which taxonomic information is most efficiently stored. The most straightforward view of taxonomies is that they are relatively permanent knowledge structures, consisting of nested hierarchies connected by class inclusion relations. However, research (Ross & Murphy, 1999; Proffitt, Coley, & Medin 2000; Sloman, 1998) often demonstrates the more fluent nature of taxonomies. People often abandon inclusion relations in favor of similarity (Sloman, 1998). Moreover, organization strategies other than taxonomic are apparently established in memory and are at times preferred over taxonomic organization (Ross & Murphy, 1999). Similarly, although basic levels have long been considered the level at which hierarchical information is most efficiently stored, research would suggest that taxonomic levels vary in status depending on level of expertise (Tanaka & Taylor, 1991), distinctiveness (Murphy & Brownell, 1985) and context (Murphy & Wisniewski, 1989).

A common thread running through some of this research (e.g., Ross & Murphy, 1999; Tanaka & Taylor, 1991) is that knowledge can affect which taxonomic level has "privileged status" (Tanaka & Taylor, 1991) and whether or not taxonomic structure is the preferred structure of conceptual organization (Ross & Murphy, 1999). It is also possible that prior knowledge interacts with taxonomic relationships in other ways. Understanding the relationship between prior knowledge and taxonomy cuts to the core of many issues currently under review in the category literature, including the psychological status of taxonomies and how levels of taxonomy may differ from one another. The next section introduces a learning paradigm intended to examine some of the key issues presented earlier in this paper. Beginning with an experiment in which knowledge is largely removed, and then introducing prior knowledge in the form of category labels this paper, this initiates an investigation into learning within taxonomic structures. The advantage here is that by using a learning paradigm issues central to the category literature can be explored systematically and as much as possible from the beginning of learning.

1.4 Current Proposal

Surprisingly, classification research has largely neglected issues surrounding category learning over hierarchical levels. What research has been performed has focused somewhat narrowly on issues of basic level superiority (Rosch, Mervis, Gray, Johnson, &

Boyes-Braem, 1976), subordinate level specificity (Murphy & Brownell, 1985), and superordinate level distinctiveness (Murphy & Brownell, 1985). Furthermore, the only study approximating a taxonomic learning paradigm (see e.g., Murphy & Smith, 1982) focused very little on learning within taxonomies, or how learning is transferred from one taxonomic level to another. Answering this question is important, as principles of learning may be found that add insight to into how taxonomies are formed. One focus of the present study is learning within taxonomies.

Experiment 1 of this paper employs artificial and abstract materials. There are several advantages to using abstract materials. For instance, perceptual and conceptual processes are deeply interconnected, making them difficult to isolate and study independently (see e.g., Goldstone & Barsalou, 1998). Using unfamiliar materials allows for some control over these processes. Furthermore, unlike descriptive studies using real-world categories (e.g., see Keil, 1989; Medin, Lynch, Coley, & Atran, 1997; Ross & Murphy, 1999), research using unfamiliar materials is more successful at revealing the mechanisms of category learning that lead to the development of conceptual structures. For instance, it is much easier to separate the influence of prior experience from experimental manipulation, which can identify the underlying learning process more clearly. Fundamentally then, use of unfamiliar materials is important for ruling out extraneous causal explanations for category learning in favor of those characterized by the category structure (e.g., see Murphy, 2002).

While the advantages of using artificial materials are apparent, some have suggested (Murphy, 2002; Schyns & Murphy, 1994) that in order to have a fuller understanding of concept formation a focus on real world contexts is crucial. They argue that while investigations of category learning using artificial materials may reflect aspects of how new categories are learned in the real world, there is some question as to what degree the categories learned are ecologically valid (this is particularly so for simple artificial materials). With this argument partly in mind, materials used in the following experiments are derived from real world objects, specifically, musical instruments.

Importantly, even though the attributes of these objects are abstract, the dimensions and their values approximate the real object. In other words, the dimensions and their values have ecological validity. However, because they are abstract, their exact relationship to the category may not be immediately apparent. For example, participants may have a clear idea of how the dimension *weight* is related to the instrument flute but they may be less clear on how the dimension *resonating frequency* relates to the instrument flute. This is particularly so in Experiment 1, where artificial labels are used in place of meaningful labels. However, in Experiment 2, prior knowledge in the form of meaningful labels that describe the object are introduced. Prior knowledge introduced in this way, means that categorization is not only brought closer to ecological validity but, according to prevailing arguments (see e.g., Murphy, 2002), there is an added advantage of facilitating category learning. One clear expectation then, is for meaningful labels to facilitate learning within taxonomic structures. After all, having prior knowledge of “wind instruments” also cues relevant attributes important to different types of “wind instruments”. The one caveat to this prediction follows from the use of abstract dimensions. As the relationship between abstract information, prior knowledge and learning within taxonomies has largely gone unexplored the outcome is difficult to predict. One possibility is that mapping of item information to the category is more difficult (see e.g., Heit & Bott, 2002; Heit et al. 2004).

Experiment 3 extended an examination of prior knowledge influences by introducing instructional information intended to boost the manipulation of knowledge associated with the meaningful label. This experiment also focused on whether prior knowledge interacts differently depending on taxonomic level. In Experiment 4, two additional items were introduced that varied in structure from that of earlier items. Prior research (see e.g., Kaplan & Murphy, 2000; Murphy & Kaplan, 2000) using thematic features has demonstrated that facilitation effects of prior knowledge increase as the number of features related to the category increase. A related question is asked in Experiment 4, but using feature-based information that is abstract. Another question in Experiment 4 centers on potential boundaries and limitations of prior knowledge learned within taxonomic structures when abstract item information is in play. This question extends a rationale suggested in Experiment 3, which suggested that slower reaction times on the part of the meaningful label group resulted from their finding abstract information inconsistent with strongly held prior experiences. In Experiment 5, items introduced in Experiment 4 were added to earlier stages of learning. Some research (Heit & Bott, 2000) has found that facilitation effects due to prior knowledge vary in magnitude over the course of learning. In that research, effects of prior knowledge manifested only after enough “data” had been collected to make an informed choice. One question in Experiment 5 is whether an opposite trend emerges when item information is abstract. Further examination of factors that were reported to have been responsible for findings in earlier experiments was also addressed.

1.5 General Method

All five experiments in this dissertation use the same materials and very similar procedures. In this section, I present information that is common across the five experiments and that is critical for understanding the purpose of the experiments. The materials were created in order to instantiate the two levels of hierarchy, which I will designate a basic level (four contrasting categories) and a superordinate level (two contrasting categories).

The individual item structures used in each of the experiments are depicted in Table 1. Importantly, participants see the same items whether they are learning the basic level categories or the superordinate level categories. Item dimensions and values used in each of the five experiments were derived from musical instruments, specifically *flute*, *saxophone*, *drum,* and *bell*. In turn, *flute* and *saxophone* belong to the superordinate category *wind* and *drum* and *bell* belong to the superordinate category *percussion.* Items presented at both basic and superordinate levels had features from contrasting categories added. For example, focusing on basic level categories depicted in Table 1, the numbers *1* on each dimension tends to indicate the category *flute*, the numbers *2* tends to indicate the category *saxophone*, the numbers *3* the category *drum*, and the number *4* the category *bell*. When considering exemplar A1 for the category *flute* (see table 1) four dimensional values (1111) indicates the category *flute*, one dimensional value (2) indicates the category s*axophone*, and one dimensional value (3) indicates the category *drum*. Twelve exemplars belong to each basic level category. When considering superordinate categories all the numbers *1* and *2* tend to indicate the *wind* category and all numbers *3* and *4* indicate the *percussion* category. Thus when focusing on Exemplar A1, 5 dimensional values (11112) indicate the category *wind* and 1 dimensional value (3) indicates the category *percussion*. Twenty-four exemplars belong to each superordinate category.

Table 1.

Abstract item structures used for basic and superordinate level categories

					Wind								
	Flute							Saxophone					
Exemplar	D1	D2	D3	D4	D5	D6	Exemplar	D1	D2	D3	D4	D5	D6
A1	1	1	1	1	2	3	B1	2	2	2	2	1	3
A2	1	1	1	1	2	4	B2	2	2	2	2	1	4
A3	1	1	1	2	3	1	B3	2	2	2	1	3	2
A4	1	1	1	2	4	1	B4	2	2	2	1	4	2
A5	1	1	2	3	1	1	B5	2	2	1	3	2	2
A6	1	1	2	4	1	1	B6	2	2	1	4	2	2
A7	1	2	3	1	1	1	B7	2	1	3	2	2	2
A8	1	2	4	1	1	1	B8	2	1	4	2	2	2
A9	2	3	1	1	1	1	B9	1	3	2	2	2	2
A10	2	4	1	1	1	1	B10	1	4	2	2	2	2
A11	3	1	1	1	1	2	B11	3	2	2	2	2	1
A12	4	1	1	1	1	2	B12	4	2	2	2	2	1
					Percussion								
	Drum							Bell					
C1	3	3	3	3	4	1	D1	4	4	4	4	3	1
C2	3	3	3	3	4	2	D2	4	4	4	4	3	2
C3	3	3	3	4	1	3	D3	4	4	4	3	1	4
C4	3	3	3	4	2	3	D4	4	4	4	3	2	4
C5	3	3	4	1	3	3	D5	4	4	3	1	4	4
C6	3	3	4	2	3	3	D6	4	4	3	2	4	4
C7	3	4	1	3	3	3	D7	4	3	1	4	4	4
C8	3	4	2	3	3	3	D8	4	3	2	4	4	4
C9	4	1	3	3	3	3	D9	3	1	4	4	4	4
C10	4	2	3	3	3	3	D10	3	2	4	4	4	4
C11	1	3	3	3	3	4	D11	1	4	4	4	4	3
C12	2	3	3	3	3	4	D12	2	4	4	4	4	3

Note. *Each exemplar for basic level (A1-D12) and superordinate level categories (A1-D24) are represented by a row in the table*

It is important to note that dimensional values for basic level categories that are in the same superordinate categories share a greater degree of similarity to one another than to basic level categories that are in a different superordinate. That is, the dimensional values for *flute* and *saxophone* are far more similar to one another than to *drum* or *bell* (and vice versa). This characteristic is common of real world categories. For example, *dogs* and *cats*, which are both *animals*, are far more similar in size to each other than either is to, say, a *house* or a *bank*, both of which are *buildings*. The dimension and values used with each dimension are presented in Table 2.

Table 2

Dimensions and values for items used in each of the 5 Experiments

Dimension	FLUTE	SAX
weight	.4kg	1kg
complexity	5p	6p
internal volume	24cu	76cu
energy required	107e	130e
resonant frequency	180db	162db
total number of possible objects	8ob	10ob
	DRUM	BELL
weight	42kg	55kg
complexity	3p	2p
internal volume	821cu	1009cu
energy required	248e	195e
resonant frequency	73db	90db
total number of possible objects	15ob	18ob

Note. Items are presented in their prototypical form.

Groups used in each of the experiments are depicted in Table 3.

Table 3

Participants randomly assigned to one of the following Experimental groups

Group	Phase 1	Phase 2
Repeated	Superordinate	Superordinate
Repeated	Basic level	Basic level
Transfer	Superordinate	Basic level
Transfer	Basic level	Superordinate

Note. *Participants in repeated groups experience same categories at phase 1 and 2. Participants in transfer group experience different categories at phase 1 and 2.*

Participants in the basic-superordinate group learned basic level categories in phase 1, and then transferred to superordinate categories in phase 2. Participants in the superordinate-basic group learned superordinate categories in phase 1, and then transferred to basic level categories in phase 2. Participants in the superordinate-superordinate group learned superordinate categories in phase 1, then repeated learning of superordinate categories in phase 2. Participants in the basic-basic group learned basic categories in phase 1 and repeated learning of basic categories in phase 2. Four categories were learned for basic levels and two categories were learned for superordinate categories. Participants who repeated taxonomic levels (the superordinate-superordinate group and the basic-basic group) learned the same items and taxonomic levels in phase 1 and 2. Participants in the taxonomic transfer groups (the superordinate-basic group and the basic superordinate-group) learned the same items in phase 1 and 2 but transferred to different taxonomic levels from phase 1 to 2.

The learning order manipulated by means of these experimental groups was designed with the intent of examining the possibility of positive taxonomic transfer (i.e. whether more efficient processing and learning of one taxonomic level follows from learning a different taxonomic level).

Chapter II

Experiment 1

The debate over the nature of taxonomic relations has been extensive (see e.g., Collins & Quillian, 1969; Murphy & Ross, 1999; Rips et al. 1973; Sloman, 1998). Questions often focused on whether taxonomies were relatively stable knowledge structures characterized by property inheritance (Collins & Quillian, 1969), strategies adopted when inferring taxonomic relations (Sloman, 1998), and how knowledge influences the use of taxonomic structures (Murphy & Ross, 1999). Using a mixture of artificial and abstract materials, this experiment initiates an investigation into the learning and use of taxonomic relations. Because participants have no prior knowledge that the information they are asked to categorize is related taxonomically, findings in favor of positive taxonomic transfer (more efficient processing and learning of one taxonomic level as a result of learning a different taxonomic level) would support the notion that people tend to naturally adopt taxonomic relations.

Two specific questions are asked in this experiment. Question 1 asks whether processing and learning following transfer to new taxonomic levels is as efficient as when levels are repeated. Evidence in support of this finding would suggest perfect taxonomic transfer. Question 2 asks whether more efficient processing and learning occurs for taxonomic transferred conditions than for the same conditions learned in phase 1. For example, is performance better for participants in the superordinate-basic phase 2 condition than for participants in the basic-superordinate phase 1 condition? This question is important for assessing whether taxonomic learning of any kind has occurred.

Finally, though outside the scope of the present research, category learning of the kind in this experiment goes to the heart of a central issue in the category literature. Specifically, one argument suggests that object learning does not require prior knowledge, and instead low-level processes like selective attention to perceptual properties can lead to the development of conceptual knowledge (Smith, 1989). Proponents of this view have argued for the interdependence of perceptual and conceptual similarity. An alternative view is illustrated by Goodman's argument (Goodman, 1992/1972) that object learning requires one to identify dimensions. Therefore, the knowledge of dimensions and beliefs about their importance should come prior to object learning. Though not directly addressed, findings in this and subsequent experiments may hint at solutions to this question and directions for future research.

2.1 Method

2.1.1 Participants

One hundred and three university undergraduates volunteered to participate in this experiment for partial course credit. Participants failing to perform beyond chance, or with average reaction times exceeding 30 seconds were removed from all analyses. A total of 15 participants were removed. Twelve participants were removed for exceeding a reaction time of 30 seconds and 3 were removed for failing to meet the learning criterion.

2.1.2 Materials

In Experiment 1 the meaningful labels described in the previous section were replaced with artificial labels. The meaningful label "flute" was replaced with the label "AAX", "saxophone" with "SSX", drum with "KKX", and "bell" with "LLX". The label "wind" was replaced with "DAX" and the label "percussion" was replaced with "JAX".

2.1.3 Procedure

All instructions and reminders appeared on a computer screen. Learning was conducted on Macintosh computers using the program Super Card.

In each group items were presented in the center of the screen with category labels situated directly above. Participants had as much time as needed to study each item. For superordinate level categories, participants indicated their category choice by pressing the *D* or *J* key (the first letter of each category label). In order to make category decisions for basic level categories participants' pressed the keys *A, S, K*, or *L*. After pressing a key a message appeared below the exemplar informing the participant of the correct answer (e.g., "the correct answer is DAX"). The answer appeared on the screen for 2 seconds so the participant could study the correct answer. Items advanced automatically after 2 seconds. Following each trial block, the participant pressed the letter *R*, which caused the next trial to start and the first item to appear on the screen. Each block of trials contained all 48 exemplars in random order. Learning continued until 4 blocks had been completed, where upon participants either repeated the same taxonomic level or transferred to a new taxonomic level for the remaining 4 blocks.

2.2 Results and Discussion

Four principle sets of comparisons were made in this experiment. First, repeated category conditions were compared with taxonomic primed conditions. Thus, the superordinate-superordinate phase 2 condition was compared with basic-superordinate phase 2 condition and the basic-basic phase 2 condition was compared with superordinate-basic phase 2 condition. Findings showing that participants in transferred categories perform as well as participants in repeated category conditions would provide support for learning of taxonomic relations between categories. Second, categories learned for the first time at phase 1 are compared with categories learned for the first time at phase 2. That is, superordinate-basic phase 1 conditions were compared with basic-superordinate phase 2 conditions, and basic-superordinate phase 1 conditions were compared with superordinate-basic phase 2 conditions. These two sets of comparisons assessed the question of taxonomic learning by asking whether there was an advantage to learning categories after taxonomic priming as opposed to learning categories without taxonomic priming. Greater number of correct responses and faster reaction time responses for phase 2 conditions would support an affect of taxonomy.

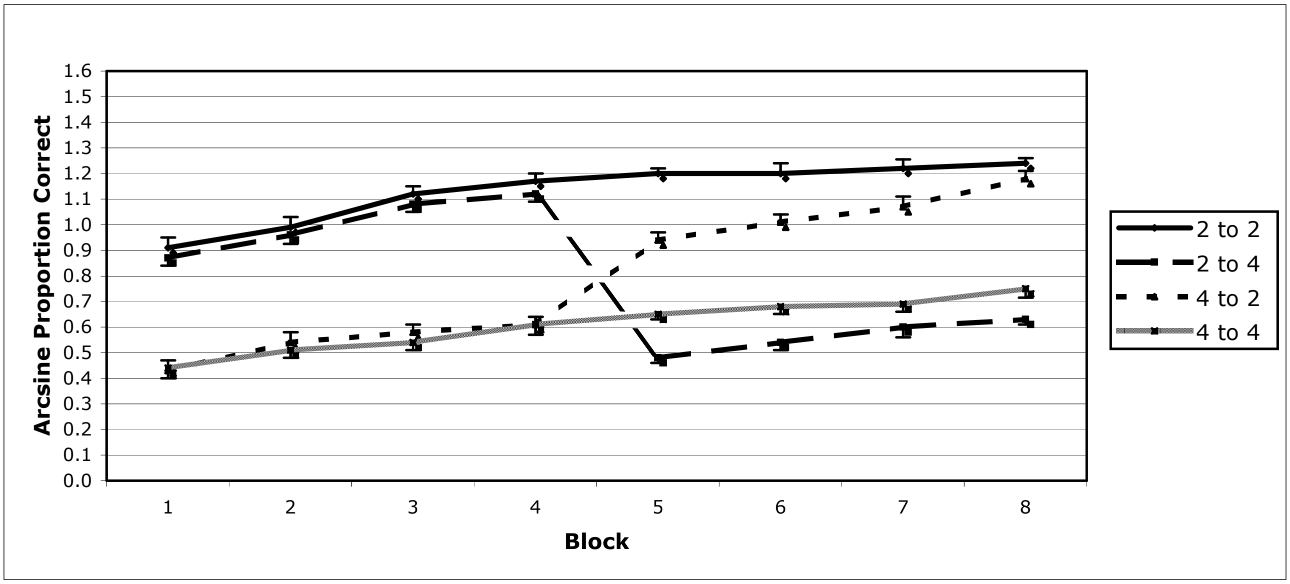
The key results are shown in Figures 1 and 2. Evidence of taxonomic learning is somewhat equivocal. Performance in phase 2 of the repeated conditions is substantially better than the taxonomic transfer condition. However, the results also seem to suggest some taxonomic learning in that performance for the basic-superordinate phase 2 condition is slightly better than the phase 1 superordinate condition. However, only minor observable differences seem apparent between the superordinate-basic phase 2 condition and the basic phase 1 condition.

2.2.1 Accuracy

For the following reasons arcsine transformations are employed in each of the five Experiments. Firstly, because proportions are bounded at zero at the low end of the scale and at one at the high end of the scale, they may not linearly relate to other continuous variables. Arcsine transforms dependent variables in the form of proportions by stretching out the tails of distributions of proportions. The arcsine transformation also has the added benefit of reducing violations of sphericity. For each of the 5 experiments two sets of analyses were performed for accuracy, one for untransformed data and for arcsine transformed data. Means for both untransformed and transformed data can be viewed in Appendix I-IV. Untransformed analyses can be viewed in Appendix V-IX. Arcsine transformed data are presented in the body of this paper.

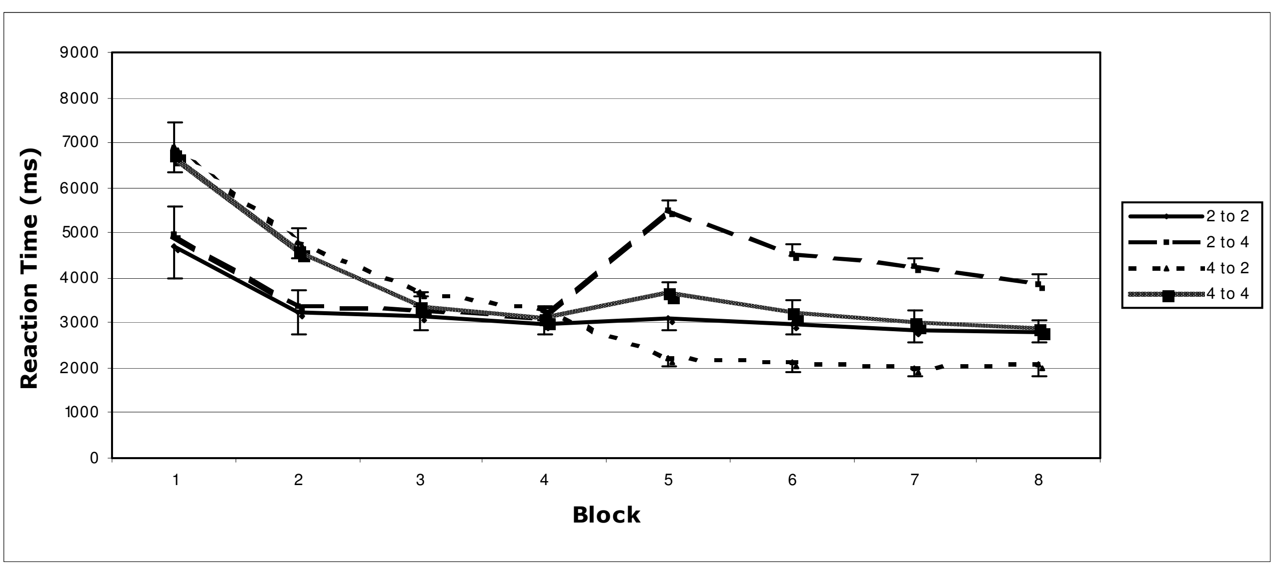
Though arcsine transformations were employed it is also important to note several observations for the untransformed data. Proportions for untransformed accuracy data were all based on the same number of observations. Furthermore, the accuracies were reasonably normal, so restriction of range usually associated with accuracy proportions was not that much of a problem. Moreover, the variances were not as different as one might expect, however they at times differed enough to violate sphericity assumptions. Thus, while arcsine transformations were employed to help bring variances closer to an assumption of equality the actual violations of the sphericity were minimal.

Figure 1. *Repeated number combinations (e.g., 2 to 2) presented in graph margins depict repeated taxonomic conditions. Mixed number combinations (e.g., 2 to 4) depict taxonomic transferred conditions.*



Item learning for taxonomic groups in Experiment 1

Figure 2. *Repeated number combinations (e.g., 2 to 2) presented in graph margins depict repeated taxonomic conditions. Mixed number combinations (e.g., 2 to 4) depict taxonomic transferred conditions.*



Item response times for taxonomic groups in Experiment 1

Mixed factorial ANOVA's were performed on each set of analyses. The first set of analyses examines the question of perfect taxonomic transfer learning. The main question is whether participants who had no prior experience with categories can perform as well as participants who have had prior experience. For example, can participants learning basic-superordinate phase 2 categories perform as well as participants learning superordinate-superordinate phase 2 categories? A finding favoring this outcome would suggest a benefit to taxonomic learning as participants who have repeated categories (superordinate-superordinate phase 2 condition) have the clear advantage of seeing the same item and category structure over participants who have seen the same items but in the presence of a different category structure (basic-superordinate phase 2 condition).

Training block refers to learning over repeated blocks (e.g., within group or condition performance), group condition refers to comparisons made between taxonomic conditions (e.g., comparisons between basic-superordinate phase 2 condition and basic-basic phase 2 condition). Participants learned four training blocks in each condition.

First, focusing just on superordinate phase 2 conditions, analysis showed a significant main effect of training block ($\underline{F}$ (3, 132) = 12.41, $\underline{p} < .001$) an interaction between training block and group ($\underline{F}$ (3, 132) = 6.12 $\underline{p} < .001$) and a main effect of group condition ($\underline{F}$ (1, 44) = 7.04, $\underline{p} < .01$). Participants in the superordinate-superordinate phase 2 condition performed better than participants in the basic-superordinate phase 2 condition on their respective first ($\underline{t}$ (44) = 24.22, $\underline{p} < .001$) second, ($\underline{t}$ (44) = 7.16, $\underline{p}$ = .01) and third ($\underline{t}$ (44) = 4.21, $\underline{p} < .05$) training blocks.

Next, focusing on basic phase 2 conditions, analysis showed a significant main effect of training block ($\underline{F}$ (3, 132) = 11.74, $\underline{p} < .001$) and a main effect of group condition (F (3, 132) = 6.12 $p < .001$). Participants in the basic-basic phase 2 condition performed better than participants in the superordinate-basic phase 2 condition on their respective first (t (44) = 27.87, $p < .001$) second (t (44) = 10.45 $p < .003$) third (t (44) = 4.17, $p < .05$) and fourth (t (44) = 6.23, $p < .02$) training blocks.

Results failed to provide support for the first question asked in this section; participants transferring to new taxonomic levels (e.g., basic-superordinate phase 2 condition) did not perform as well as participants who repeated learning of taxonomic levels (e.g., superordinate-superordinate phase 2 condition).

The next set of analyses examines the question of whether any taxonomic learning occurred at all. Here all comparisons involve first time category exposures. The general idea is that if performance for participants learning phase 2 categories is superior to that of participants learning phase 1 categories there is evidence of taxonomic learning. That is, because participants in both conditions are learning particular categories for the first time, findings favoring taxonomic primed groups would suggest a learning advantage, due to experience with the taxonomically related category. This advantage may result from learning of class inclusion relations or any number of other factors. Given that participants are learning all categories for the first time comparisons are on some level standardized. However, it is important to note that participants learning the phase 2 categories have had prior exposure to items (i.e., the same items are presented in phase 1 and 2) thus may on some level have an advantage over phase 1 participants. Because there were no meaningful differences between superordinate-superordinate and superordinate-basic or between the basic-superordinate and basic-basic groups in phase 1, these pairs were combined for analysis. That is, the superordinate-basic phase 1 condition and the superordinate-superordinate phase 1 condition were combined as were the basic-superordinate phase 1 condition and the basic-basic phase 1 condition. Analyses therefore consisted of comparing the basic-superordinate phase 2 condition with the superordinate phase 1 combined condition and the superordinate-basic phase 2 condition with basic phase 1 combined condition.

First, focusing just on superordinate conditions, analysis showed a significant main effect of training block (F (3, 132) = 37.05, $p < .001$). No main effect of group condition was found; participants learning items for the first time in the basic-superordinate phase 2 condition failed (F (1, 44) = 1.92, $p > .17$) to outperform participants learning items for the first time in the superordinate phase 1 condition. The interaction between group condition and training block was also statistically non-significant (F (3, 132) = .43, $p > .73$).

Next, focusing on basic conditions, analysis showed a significant main effect of training block (F (3, 132) = 25.09, $p < .001$). No main effect of group condition was found, participants learning items for the first time in the superordinate-basic phase 2 condition failed (F (3, 132) = .64, $p > .43$) to outperform participants learning items for the first time in the basic phase 1 condition. The interaction between group condition and training block was also statistically non-significant (F (1, 44) = .25, $p > .86$)

Analyses failed to provide support for the second question asked in this experiment; participants learning categories for the first time in phase 2 did not outperform participants learning categories for the first time in phase 1. Thus, these comparisons did not support effects of taxonomic learning. In the next section reaction times are examined. Questions and expected outcomes for reaction times are identical to those for accuracy.

2.2.2 Reaction times

The reaction times were averaged and submitted to mixed factorial ANOVA's, after discarding any times greater than 30 seconds. Only correct responses were analyzed. Analyses and comparisons are identical to those for accuracy. The first set of analyses examined the question of perfect taxonomic transfer by comparing first time category learning experiences with repeated category learning experiences. Effects of perfect taxonomic learning would show that participants learning categories for the first time in phase 2 perform as well as participants repeated the same categories in phase 2.

First, focusing on superordinate conditions, analysis showed a significant main effect of training block (F (3, 132) = 7.17, $p < .01$). The interaction between training block and group condition was statistically non significant (F (3, 132) = 1.18, $p > .14$) as was the main effect of group (F (1, 44) = .15, $p > .93$). No reaction time differences were evident between the basic-superordinate phase 2 condition and the superordinate-superordinate phase 2 condition.

Next, focusing on basic conditions, analysis showed a significant main effect of training block (F (3, 132) = 15.31, $p < .001$) and a main effect of group condition (F (1, 44) = 5.10, $p < .002$). Participants in the basic-basic phase 2 condition responded faster than participants in the superordinate-basic phase 2 condition on their respective first (t (44) = 3.51, $p < .001$) second (t (37) = 2.23, $p < .03$) third (t (44) = 2.26, $p < .03$) and fourth (t (44) = 2.52, $p < .02$).

Results showed partial support for effects of taxonomy in that participants who transferred from basic to superordinate categories processed items as fast as participants who transferred from superordinate to superordinate categories. Thus, participants repeating items in the presence of different taxonomic category performed as well as participants repeating items in presence of the same taxonomic category (indeed, though statistically non-significant they responded faster). However, this finding was apparent only for superordinate categories. Participants repeating basic level categories were faster than participants who transferred from superordinate to basic categories on all four blocks.

The next sets of analyses investigate whether there is any evidence of taxonomic learning. First, focusing just on superordinate conditions, analysis showed a significant main effect of training block (F (3, 132) = 7.82, $p < .001$), an interaction between training block and group condition (F (3, 132) = 5.10, $p < .002$) and a main effect of group condition (F (1, 44) = 10.11, $p < .003$). Participants learning items for the first time in the basic-superordinate phase 2 condition were faster processing items than participants learning items for the first time in the superordinate-basic phase 1 condition on their respective first (t (44) = 3.93, $p < .001$) second (t (44) = 2.17, $p < .04$) and third training blocks (t (44) = 2.52, $p < .02$).

Second, focusing just on basic conditions, analysis showed a significant main effect of training block (F (3, 132) = 7.82, $p < .001$) and an interaction between training block and group condition (F (3, 132) = 5.10, $p < .002$). Participants learning items for the first time in the superordinate-basic phase 2 condition were faster processing items than participants learning items for the first time in the basic-superordinate phase 1 condition on their respective first training block (t (44) = 2.24, $p < .03$).

When considering reaction times, the results generally favored the idea that learning categories for the first time following taxonomic transfer has advantages over learning items for the first time in phase 1, though it is important to note that processing advantages were greater for the basic-superordinate group than for the superordinate-basic group. The basic-superordinate group outperformed the superordinate-basic phase 1 condition on 3 blocks of training, whereas the superordinate-basic phase 2 condition outperformed the basic phase 1 condition on only the first block of training. As discussed next these differences in performance may follow in part from advantages to learning basic level categories first.

Two main questions were asked in this experiment. The first question asked whether participants learning a taxonomic level for the first time in phase 2 following having learned a different taxonomic level in phase 1 would perform as well as participants repeating taxonomic level in phase 2. Evidence in support of perfect taxonomic transfer effects of this kind was evident only in the form of reaction times. Moreover, this outcome was found only between the basic-superordinate phase 2 condition and superordinate-superordinate phase 2 condition. This is very weak evidence, given that accuracy data was quite strongly in the opposite direction (i.e., indicating much worse performance in the taxonomic transfer conditions). The second question asked whether participants learning a taxonomic level for the first time in phase 2 (after having learned a different taxonomic level) would outperform participants who had learned that taxonomic level in phase 1. Evidence in support of a transfer effect of this kind was also only in the form of reaction times. Participants in basic-superordinate phase 2 conditions were faster processing items than participants in the superordinate-basic phase 1 condition over the first three blocks of training, and participants in the superordinate-basic phase 2 condition were faster processing items than participants in the basic superordinate phase 1 condition on the first block of training. However, without evidence in the form of accuracy to differentiate these comparisons an affect of taxonomy is difficult to conclude. This is because it is difficult to know whether participants are responding faster in phase 2 condition as result of being primed by a different taxonomic structures or simply because they are more familiar with the items.

Interestingly, as suggested above reaction time effects were stronger for participants transferring from basic to superordinate categories than for participants transferring from superordinate to basic levels. This would suggest the simple familiarity is not the sole explanation for present findings. Another reason may follow from exposure to basic levels categories in phase 1. As noted previously basic level categories that are in the same superordinate share a greater degree of similarity to one another than basic level categories that are in different superordinates. That is, AAX and SSX have dimensional values that are similar to one another but not so similar to KKX and LLX (and vice versa). When participants were exposed to basic level categories they were in a far better position to learn structural characteristics than participants exposed to superordinate levels. Participants learning basic categories have an opportunity to compare and contrast superordinate instantiations. That is, they were not only able to compare how AAX (flute) and SSX (saxophone) were similar and different to one another but the were also able to compare how they are similar and different to KKX (drum) and LLX (bell). Conversely, participants learning superordinate categories were only able to compare two categories. As a result they did not have foreknowledge of all basic level categories and had to consider how the dimensional values transferred to additional categories. In sum, participants exposed to basic level categories had a better understanding of how dimensional values belonged to individual categories and as such were better able to generalize items to transferred categories.

In the next experiment meaningful labels are introduced. These labels identify the categories and are expected to boost detection of the taxonomic relationships.

Chapter III

Experiment 2

In Experiment 1, the materials limited access to prior knowledge and experiences. In the present experiment, identifying our items with meaningful category labels is intended to cue prior knowledge. Prior knowledge associated with category labels may play an important role in how features are interpreted and may also lead one to selectively attend to those features that are particularly relevant. For example, being informed that an item is a flute may cue both semantic information and episodic experiences a person has previously had with flutes. This info to experience can then be used to guide feature selection and categorization. Labels also imply probabilistic information about the features of the referred object (Anderson, 1991). For example, knowing that the object being categorized is a flute allows one to make an informed guess as to its weight. Prior knowledge in the form of category labels has also been shown to guide learning by providing an explanation for the properties and structure of categories (Kaplan & Murphy, 2000). Meaningful labels may be an important factor in learning the category membership of an object.

While the usefulness of meaningful labels has been demonstrated in many contexts, to the author's knowledge, how meaningful labels, real world taxonomic categories, and abstract structures interact to produce learning has largely been ignored. Much prior taxonomic research (see e.g., Markman & Callanan, 1984; Osborne & Calhoun, 1998) has focused on novel categories and labels without explicitly addressing the influence of prior knowledge cued by meaningful labels. The failure to use familiar taxonomic labels may on some level explain why prior research often demonstrates negative taxonomic findings (see e.g., Osborne & Calhoun, 1998). When labels are not strongly connected to prior knowledge, activation of feature relations is likely reduced. That is, people may not easily see how one feature is related to another. Moreover, meaningful labels denote the category and the taxonomic level, thus boosting the transfer between taxonomic levels (see below). Finally, a focus on artificial categories without reference to natural categories is often at the expense of ecological validity (see e.g., Murphy, 2002). Introducing taxonomic labels that link strongly to prior experiences is one step toward achieving ecological validity. While both meaningful labels and attributes have real world validity, the attributes used in present studies are still somewhat abstract and in the present context potentially unfamiliar. Thus, the primary focus in this experiment is on how knowledge cued by category label affects learning of taxonomic relations. Other potential influences that may affect learning of taxonomic relations can be addressed at a later date.

Based on prior research there is reason to believe that meaningful labels (see e.g., Anderson, 1991; Kaplan & Murphy, 2000) will grant a powerful means for inferring taxonomic relationships. Knowing the object being categorized is a flute should cue the categorizer to the fact that it is also a wind instrument. Furthermore, having knowledge of the object's identity allows the categorizer to infer characteristics (e.g., weight, shape, and size) central to category membership. Findings favoring taxonomic transfer would at the very least suggest that people readily access pre-stored hierarchical structures.

In sum, meaningful labels are introduced in this experiment with the expectation that they will boost detection of taxonomic relations. The main question then is whether meaningful labels result in perfect taxonomic transfer or any taxonomic transfer at all.

3.1 Method

3.1.1 Participants

One hundred and six university undergraduates volunteered to participate in this experiment for partial course credit. Participants failing to achieve performance beyond chance, or average reaction times exceeding 30 seconds were removed from analyses. In total 17 participants were removed from analyses for failing to meet learning criterion. Seven participants were removed for performing below chance, and the remaining ten were removed for exceeding reaction times of 30 seconds.

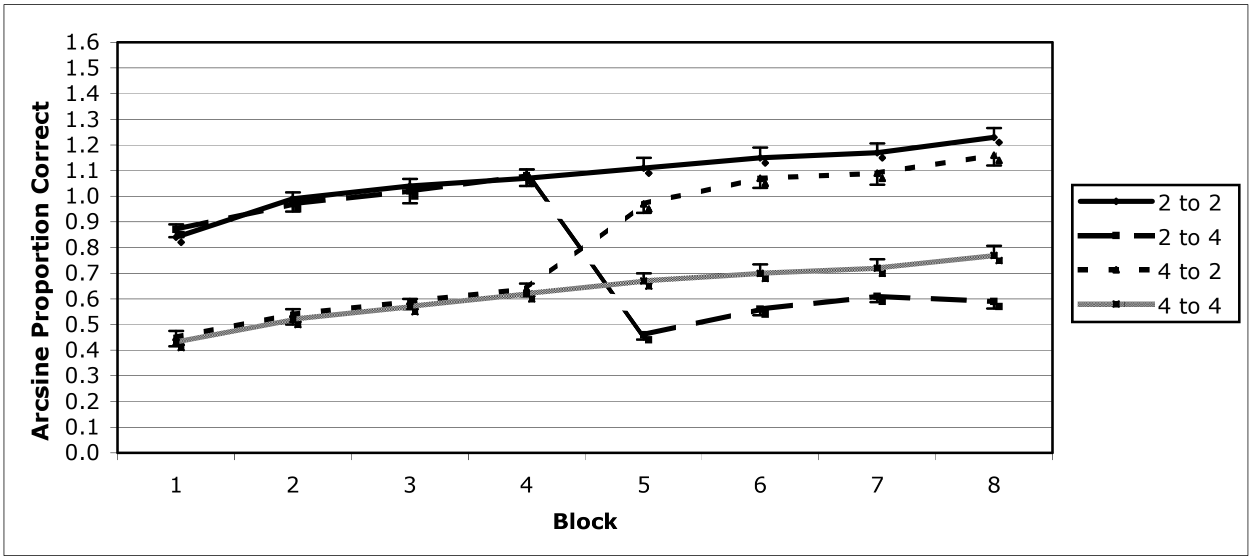
3.1.2 Materials and Procedure

With the exception of replacing artificial labels with meaningful labels that identified the categories, the materials and procedure were identical to that of Experiment 1. For the basic level category the artificial labels of AAX, SSX, KKX, and LLX were replaced with the meaningful labels of FLUTE, SAXOPHONE, DRUM and BELL respectively. For superordinate level categories the labels of WIND and PERCUSSION replaced the artificial labels DAX and JAX.

3.2 Results and Discussion

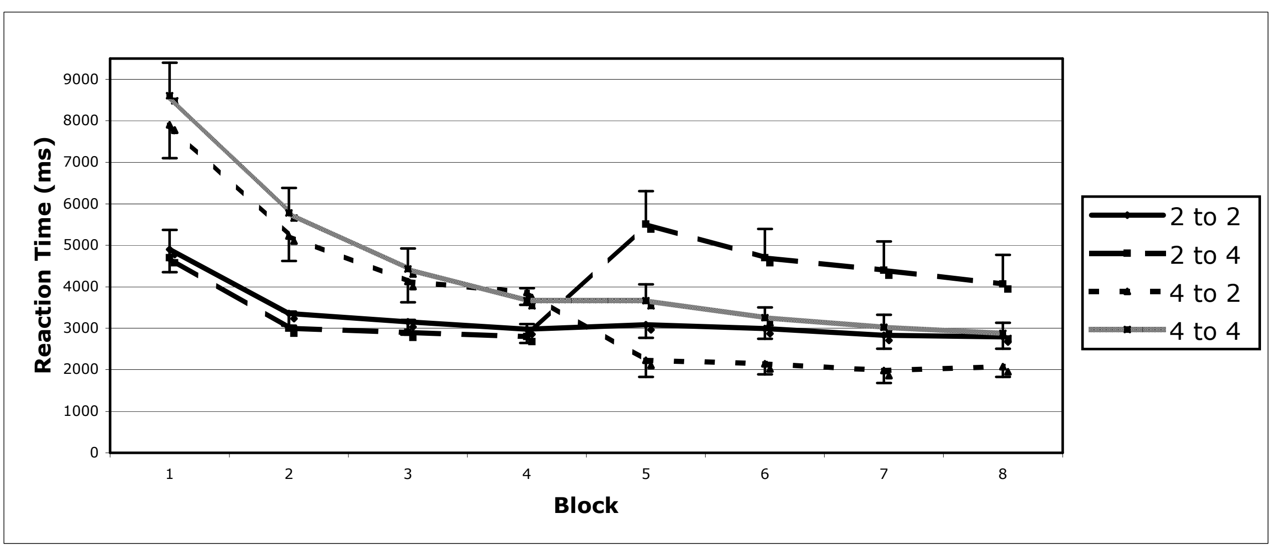
The key results are shown in Figures 3 and 4. As in experiment 1, evidence of taxonomic learning is somewhat ambiguous. Performance in the phase 2 repeated condition is substantially better than the taxonomic transfer conditions. However, the results also seem to suggest some taxonomic learning in that performance for the basic-superordinate phase 2 condition is slightly better than the phase 1 superordinate condition. Some differences are evident between the superordinate-basic phase 2 condition and the basic phase 1 condition.

Figure 3. Repeated number combinations (e.g., 2 to 2) presented in graph margins depict repeated taxonomic conditions. Mixed number combinations (e.g., 2 to 4) depict taxonomic transferred conditions.



Item learning for taxonomic groups in Experiment 2

Figure 4. *Repeated number combinations (e.g., 2 to 2) presented in graph margins depict repeated taxonomic conditions. Mixed number combinations (e.g., 2 to 4) depict taxonomic transferred conditions.*



Item response times for taxonomic groups in Experiment 2

3.2.1 Accuracy

Mixed factorial ANOVA's were performed on each set of analyses. The first set of analyses examines the question of perfect taxonomic transfer learning. The main question here is whether meaningful labels will facilitate perfect transfer learning. If so then one would expect participants in the basic-superordinate phase 2 condition to perform as well as participants in the superordinate-superordinate phase 2 condition, and for participants in the superordinate-basic phase 2 condition to perform as well as participants in the basic-basic phase 2 condition.

First, focusing just on superordinate phase 2 conditions, analysis showed a significant main effect of training block ($\underline{F}$ (3, 126) = 10.95, $\underline{p} < .001$). The interaction between group condition and training block was statistically non-significant ($\underline{F}$ (3, 126) = .38, $\underline{p} > .80$), as was the main effect of group condition ($\underline{F}$ (1, 42) = 3.18 $\underline{p} < .08$).

Next, focusing on basic phase 2 conditions, analysis showed a significant main effect of training block ($\underline{F}$ (3, 132) = 11.74, $\underline{p} < .001$) and a main effect of group condition ($\underline{F}$ (3, 132) = 6.12 $\underline{p} < .001$). Participants in the basic-basic phase 2 condition performed better than participants in the superordinate-basic phase 2 condition on their respective first ($\underline{t}$ (43) = 32.74, $\underline{p} < .001$) second ($\underline{t}$ (43) = 14.48, $\underline{p} < .001$) third ($\underline{t}$ (43) = 8.12, $\underline{p} < .01$) and fourth ($\underline{t}$ (43) = 14.36, $\underline{p} < .001$) training blocks.

Results showed partial support for perfect taxonomic transfer effect in that participants in the basic-superordinate phase 2 condition performed nearly as well participants in superordinate-superordinate phase 2 condition. As discussed shortly this finding may reflect the generalization effect explored in Experiment 1.

The next set of analyses examines the question of whether any taxonomic learning occurred. Here all comparisons involve first time category exposures. If performance for participants learning phase 2 categories is superior to that of participants learning phase 1 category there is evidence of taxonomic learning. Because there were no meaningful differences between superordinate-superordinate and superordinate-basic or between the basic-superordinate and basic-basic groups at phase 1, these pairs were combined for analysis.

First, focusing just on superordinate conditions, analysis showed a significant main effect of training block (F (3, 132) = 37.05, $p < .001$), and a main effect of group condition (F (1, 42) = 6.42, $p < .02$). Participants in the basic-superordinate phase 2 made a greater number of correct responses than participants in the superordinate phase 1 condition on their respective first (t (42) = 13.72, $p < .001$) and second (t (42) = 4.91, $p < .03$) training blocks.

Next, focusing on basic conditions, analysis showed a significant main effect of training block (F (3, 129) = 25.09, $p < .001$). The interaction between group condition and training block was statistically non-significant (F (3, 129) = 2.66, $p = .06$) as was the main effect of group condition (F (1, 43) = .03, $p > .86$). Participants learning items for the first time in the superordinate-basic phase 2 condition failed to outperform participants learning items for the first time in the basic phase 1 condition.

Analyses provided partial support for taxonomic transfer effect in that the basic-superordinate phase 2 condition outperformed the superordinate phase 1 condition. An effect of meaningful label may also be evident, as this finding was not found in Experiment 1.

3.2.2 Reaction times

The reaction times were averaged and submitted to mixed factorial ANOVA's, after discarding any times greater than 30 seconds. Only correct responses were analyzed. Analyses and comparisons are identical to those for accuracy. The first set of analyses examined the affect of meaningful labels on perfect taxonomic transfer.

First, focusing on superordinate conditions, analysis showed a significant main effect of training block (F (3, 126) = 4.43, p < .01). The interaction between training block and group condition was statistically non-significant (F (3, 126) = .51, p > .68) as was the main effect of group (F (1, 42) = .16, p > .69). Thus, no reaction time differences were evident between the basic-superordinate phase 2 condition and the superordinate-superordinate phase 2 condition.

Next, focusing on basic conditions, analysis showed a significant main effect of training block (F (3, 129) = 11.94, p < .003), as well as an interaction between group condition and training block (F (3, 129) = 4.98, p < .003) and a main effect of group condition (F (1, 42) = 7.67, p < .01). Participants in the basic-basic phase 2 condition responded faster than participants in the superordinate-basic phase 2 condition on their respective first (t (43) = 3.53, p < .001) and second (t (43) = 2.20, p < .001) training blocks.

Results showed partial support for perfect taxonomic transfer in that participants who transferred from basic to superordinate categories processed items as fast as participants who transferred from superordinate to superordinate categories. Thus, participants repeating items in the presence of different taxonomic category performed as well as participants repeating items in the presence of the same taxonomic category.

However, this finding was apparent only for superordinate categories; participants repeating basic level categories were faster than participants who transferred from superordinate to basic categories on all four blocks. These findings replicate those found in Experiment 1.

The next analyses examine the possibility of any taxonomic transfer by comparing taxonomic transfer condition phase two performance to the performance in the matched conditions at phase 1. First, focusing just on superordinate conditions, analysis showed a significant main effect of training block (F (3, 132) = 7.82, $p < .001$) an interaction between training block and group condition (F (3, 132) = 5.10, $p < .002$) and a main effect of group condition (F (1, 44) = 10.11, $p < .003$). Participants learning items for the first time in the basic-superordinate phase 2 condition processed items faster than participants learning items for the first time in the superordinate phase 1 condition on their respective first (t (42) = 3.93, $p < .001$) second (t (42) = 2.17, $p < .04$) and third training blocks (t (42) = 2.04, $p < .04$).

Second, focusing just on basic conditions, analysis showed a significant main effect of training block (F (3, 129) = 35.74, $p < .001$). The interaction between training block and group condition was statistically non-significant (F (3, 129) = 1.29, $p > .31$) as was the main effect of group condition (F (1, 43) = .02, $p > .88$).

Reactions time results in this Experiment were very similar to those found in Experiment 1. Results generally favored the idea that learning taxonomically transferred categories for the first time in phase 2 had advantages over learning the same categories for the first time in phase 1. Moreover processing advantages were greater for the basic-superordinate group than for the superordinate-basic group.

Results in this Experiment were on the whole similar to those found in Experiment 1. Taxonomic transfer effects were generally in the form of reaction times. Moreover, the reaction time effects generally supported the idea of generalization. Phase 2 learning following exposure to basic level categories in phase 1 was overall faster than phase 2 learning following exposure to superordinate categories in phase 1. Indeed, the one finding in this experiment differing from that of the previous experiment also seems to support generalization. Participants in basic-superordinate phase 2 conditions produced more correct responses than participants learning superordinate categories in phase 1. However, no advantages related to correct responses were apparent for phase 2 conditions following having learned superordinate categories in phase 1. Overall these findings correlate well with the idea that learning basic level categories facilitate generalization to new (superordinate) categories.

Importantly, taxonomic transfer effects in the form of correct responses were found in this Experiment. This suggests that meaning attached to the label may have contributed to the transfer effect. However, without direct comparisons between the meaningful and artificial label groups this observation is difficult to draw with certainty. Moreover, only one taxonomic transfer effect in the form of correct responses was found. Thus, it is difficult to draw a final conclusion about the overall affect of taxonomic transfer. Indeed, the only real conclusions that can be drawn are that meaningful labels may have facilitated generalization, and that reaction times in this experiment are slightly higher than that of Experiment 1.

Given prior evidence demonstrating facilitation effects of prior knowledge (see e.g., Murphy 2002), the failure to find clear effects of taxonomy in this experiment is surprising. The expectation was for the category label to clarify taxonomic relations. One reason for this pattern may result from some dimensions in this experiment being relatively abstract (at least within the present context). Participants may have had difficulty drawing a clear parallel between the instrument and the dimension (e.g., may not have had a clear idea of how resonating frequency, internal volume etc. related to the instrument). Another reason for failing to see strong prior knowledge effects is that in comparison to prior research (see e.g., Kaplin and Murphy, 2000; Spalding and Murphy, 1996) the dimensions and feature relations in this experiment are relatively weak.

In the next experiment instructions are introduced which clarify qualities of features and how one feature relates to another. Moreover, effects for both artificial and meaningful label groups are examined in a single experiment.

Chapter IV

Experiment 3

The previous experiments in this paper explored artificial and meaningful label groups independently, making direct comparisons between these groups problematic. In this experiment both artificial and meaningful label groups are included. Additionally, knowledge of feature relations is introduced. This information makes clear connections between instrument features. For example, instruments having greater weight will on average have lower tones. Cueing knowledge of this kind is intended not only to strongly define how one feature is related to another, but also to clarify taxonomic relations. The expectation is that in comparison to when the taxonomic label is unknown, knowledge of feature relations will boost the manipulation of knowledge associated with the meaningful label. On the other hand, such relations should not be particularly meaningful when there are no meaningful labels to which the participants can attach the relations. That is, the relations are highly abstract, and therefore may have little impact on their own.

The central question in this experiment is whether taxonomic relations are easier to learn in the presence of meaningful or artificial labels. One clear expectation is that meaningfully named categories are easier to acquire than abstract named categories because knowing the name of the object activates prior knowledge and experiences associated with that object (see e.g., Murphy, 2002; Waxman & Markow, 1995). Furthermore, participants exposed to meaningful labels in this experiment have knowledge of both instrument types and feature relations, while those exposed to artificial labels have knowledge of feature relations alone. Nevertheless, because no demonstrable benefit was found when using meaningful labels in Experiment 2, it is important to consider counter arguments. Firstly, while exposure to meaningful labels results in having knowledge of instruments, the features themselves are abstract. Even given knowledge of feature relations, it is unclear how abstract feature-based information will interact with deeply embedded prior experiences. Some research suggests (e.g., Heit, 2000) that performance is negatively affected under these kinds of conditions. Moreover, people often have only superficial knowledge of an object (see e.g., Keil, 2003). Thus, exposure to dimensions involving deeper object knowledge like resonating frequency and internal volume may be outside the scope of participants' prior experiences for musical instruments. As a result, participants may have difficulty reconciling unfamiliar information with their expectations. In contrast, participants exposed to artificial labels are unlikely to consider the relationship between label and attributes. Thus, mapping features with the label may be easier when both are abstract.

In sum, the main question asked in this experiment is whether introducing knowledge of feature relations will boost the manipulation of knowledge associated with the meaningful label. Whether a facilitation affect of prior knowledge is found may in part depend on how well item information maps onto to prior experiences.

4.1 Method

4.1.1 Participants

One hundred and eighty three university undergraduates volunteered to participate in this experiment for partial course credit. In total, 19 participants were removed from analyses for failing to meet learning criterion. Six participants were removed for performing below chance, with the remaining 13 removed for exceeding reaction times of 30 seconds.

4.1.2 Materials and Design.

This experiment used the same materials as in the previous experiments, except that in the current experiment, additional instructions were introduced. These instructions informed participants how some attributes were related. For example participants in both groups were informed that instruments light in weight generally played at higher frequencies whereas heavier instruments generally played at lower frequencies.

4.1.3 Procedure

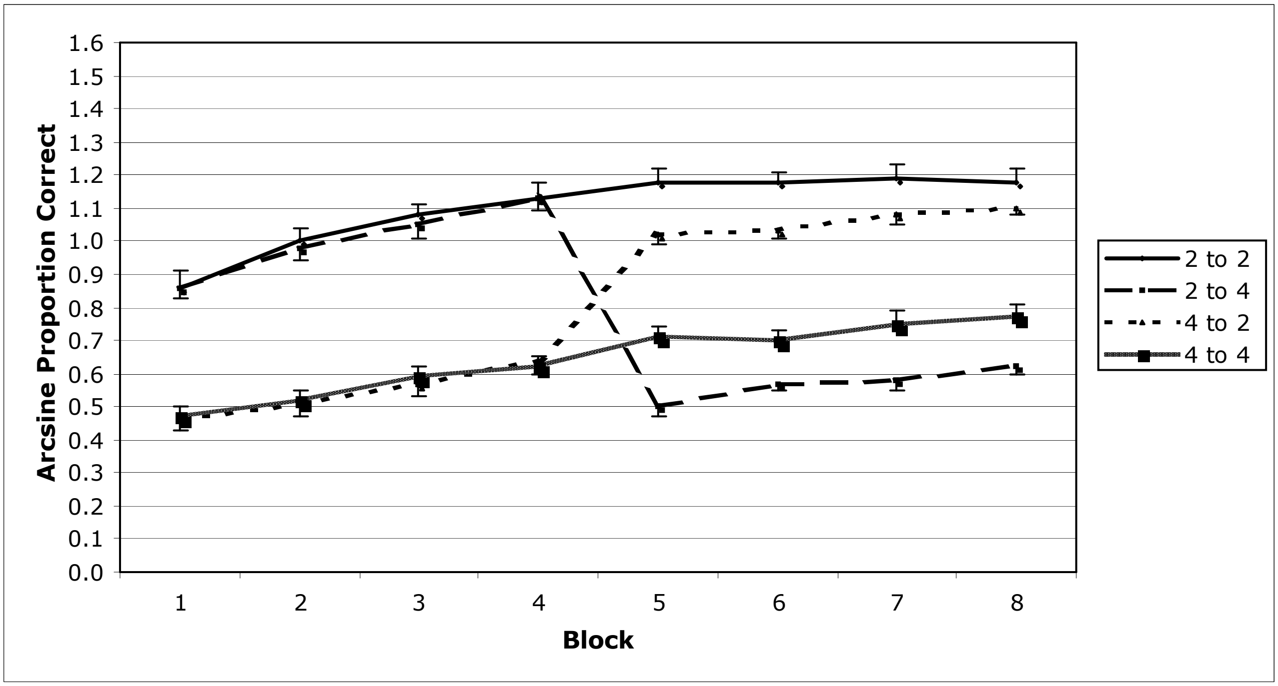
Experiment 3 procedures replicate those of Experiment 1 and 2 with one change. Before beginning category-learning participants in each condition read instructions introducing them to relations between features. Participants in the meaningful label group were provided with real instrument names, while participants in the artificial label group were given the meaningless label. In neither case was any direct connection made between the labels and the features.

4.2 Results and Discussion

Note that for this and subsequent experiments individual group taxonomic performance will not be examined. Clear taxonomic transfer effects were not found in Experiments 1 and 2. The main finding found in those experiments was that phase 2 learning following exposure to basic level categories in phase 1 was overall faster than Phase 2 learning following exposure to superordinate categories in phase 1. In general, the current experiment shows the same pattern of minimal taxonomic transfer. Taxonomic transfer is found only in reaction times and in only in transfer from basic to superordinate. These results have been established in previous experiments and remained constant over the next 3 experiments. These analyses are available upon request.

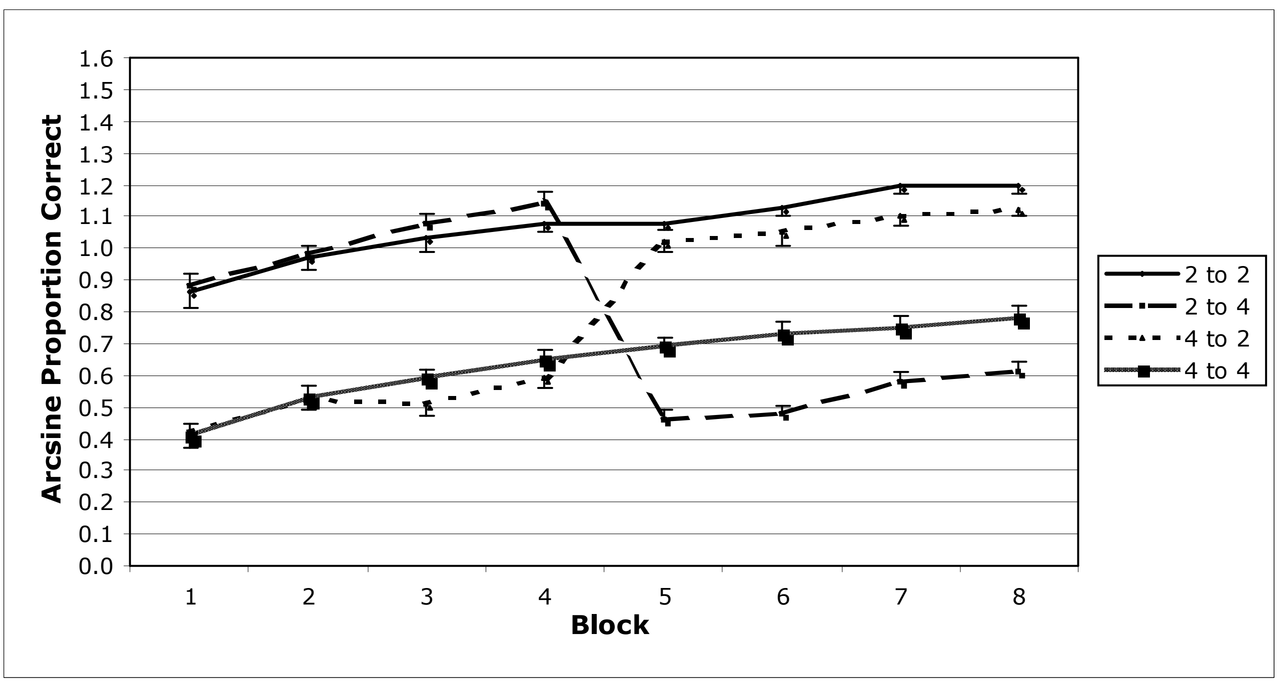
The key results are shown in Figures 5, 6, 7, and 8 separately for the artificial and meaningful groups. First, focusing on accuracy ratings (see Figures 5 and 6) it seems clear that groups did not differ on correct responses. Instructions failed to boost meaningful content associated with the label. Next focusing on reaction times (see Figures 7 and 8) it is also clear that performance for the artificial group is faster, particularly for basic level categories. This result is opposite that predicted if prior knowledge had facilitated learning and consistent with the idea that item information interferes with prior experiences.

Figure 5.



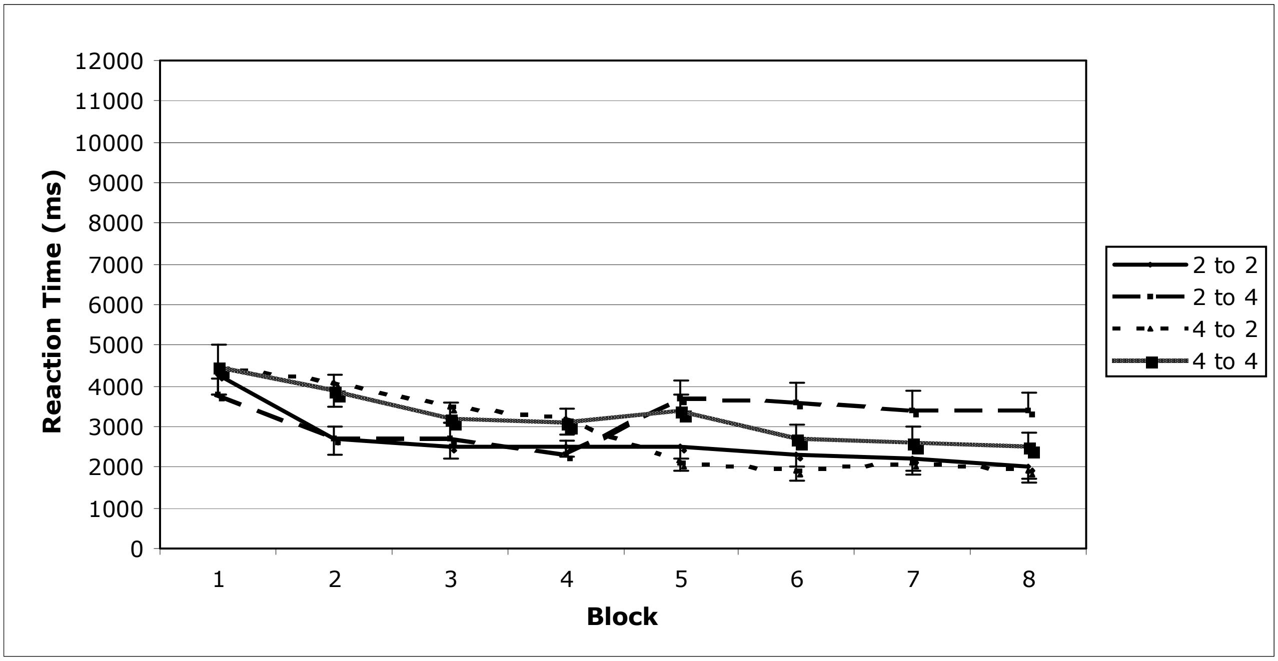
Item learning for artificial group in Experiment 3

Figure 6.



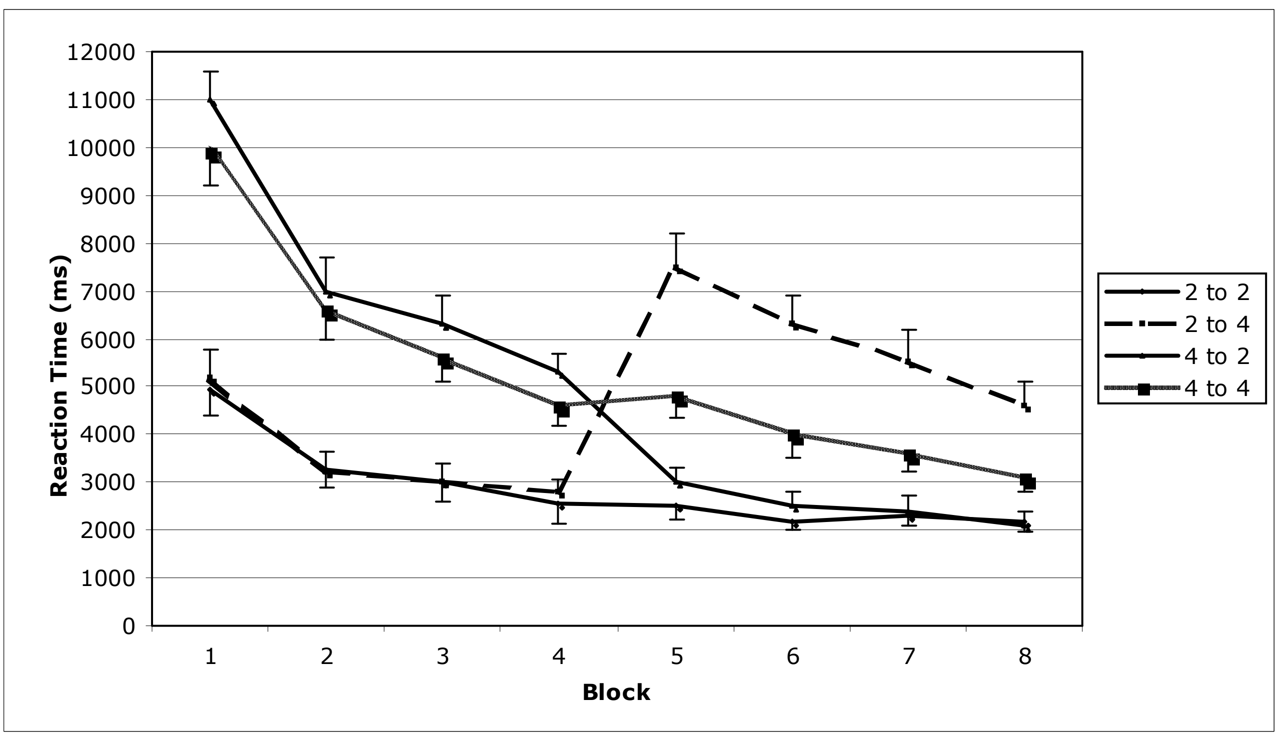
Item learning for meaningful group in Experiment 3

Figure 7.



Response times for artificial group in Experiment 3

Figure 8.



Response times for meaningful group in Experiment 3

4.2.1 Accuracy

The first set of analyses focuses on taxonomic transfer effects. Examination of taxonomic transfer effects in the present experiment focuses solely on comparisons between the artificial and meaningful groups in the basic-superordinate phase 2 conditions and superordinate-basic phase 2 conditions. These comparisons show whether one group benefits more than the other from taxonomic priming. A finding favoring the meaningful group would also suggest that instructions boosted the manipulation of knowledge associated with the meaningful label.

First when comparing mean transfer differences between the meaningful basic-superordinate phase 2 condition and the artificial basic superordinate phase 2 condition results showed a significant main effect of training block (F (3, 243) = 28.66, $p < .001$). The interaction between group condition and training block was statistically non-significant (F (3, 243) = .02, $p > .80$) as was the main effect of group (F (1, 81) = 1.35, $p > .25$).

Next, when comparing mean transfer between meaningful superordinate-basic phase 2 condition and the artificial superordinate-basic phase 2 condition results showed a significant main effect of training block ($\underline{F}$ (3, 243) = 44.60, $\underline{p} < .001$). The interaction between group condition and training block was statistically non-significant ($\underline{F}$ (3, 243) = .18, $\underline{p} > .91$) as was the main effect of group ($\underline{F}$ (1, 81) = 2.90, $\underline{p} > .09$).

In sum, no differences were found between meaningful and artificial label groups following transfer to a new taxonomic level.

The next set of comparisons examine phase 1 and phase 2 differences between artificial and meaningful groups. These comparisons include superordinate and basic level phase 1 conditions, as well as superordinate and basic level phase 2 conditions. These comparisons are important for further examining whether instructions facilitate learning for the meaningful label group. Findings showing better performance for the meaningful group would support this idea.

First, focusing on superordinate phase 1 conditions for artificial and meaningful groups, results showed a significant main effect of training block ($\underline{F}$ (3, 117) = 56.73, $\underline{p}$ < .001). The interaction between group condition and training block was statistically non-significant ($\underline{F}$ (3, 117) = 1.95, $\underline{p}$ > .12) as was the main effect of group ($\underline{F}$ (1, 39) = 1.40, $\underline{p}$ > .24).

Second, focusing on basic phase 1 conditions for artificial and meaningful groups, results showed a significant main effect of training block ($\underline{F}$ (3, 117) = 111.24, $\underline{p}$ < .001). The interaction between group condition and training block was statistically non-significant ($\underline{F}$ (3, 117) = .90, $\underline{p}$ > .44) as was the main effect of group ($\underline{F}$ (1, 39) = 1.32, $\underline{p}$ > .26).

Third, comparisons between the artificial superordinate-superordinate phase 2 condition with the meaningful superordinate-superordinate phase 2 conditions showed a significant main effect of training block ($\underline{F}$ (3, 117) = 6.29, $\underline{p}$ < .02). The interaction between group condition and training block was statistically non-significant ($\underline{F}$ (3, 117) = .56, $\underline{p}$ > .65) as was the main effect of group ($\underline{F}$ (1, 39) = 2.24, $\underline{p}$ > .14).

Finally, when comparing the artificial basic-basic phase 2 condition with the meaningful basic-basic phase 2 condition, results showed a significant main effect of training block ($\underline{F}$ (3, 117) = 4.90, $\underline{p}$ < 003). The interaction between group condition and training block was statistically non-significant (F (3, 117) = .65, p > .58) as was the main effect of group condition (F (1, 39) = .21, p > .65).

In sum, analyses revealed no differences between artificial and meaningful groups for correct responses. Instructions did not have the expected effect of boosting knowledge effects associated with the meaningful labels.

4.2.2 Reaction Times

The reaction times were averaged and submitted to mixed factorial ANOVA's, after discarding any times greater than 3 seconds. Only correct responses were analyzed. It is important to keep in mind that participants learn in total 48 items for each block, 12 for each category.

The first sets of comparisons examine taxonomic transfer differences between the artificial and meaningful label groups for basic-superordinate phase 2 conditions and superordinate-basic phase 2 conditions. Expectations for these comparisons are identical to those for correct responses.

First, when comparing the meaningful basic-superordinate phase 2 and artificial basic superordinate phase 2 condition results showed a significant main effect of training block (F (3, 243) = 23.39, p < .001). The interaction between group condition and training block was statistically significant (F (3, 243) = 3.85, p < .01). Participants in the artificial basic-superordinate condition were faster processing items at block 1 (t (81) = 3.85, p = 6.11, p < .02).

Next, comparisons between the meaningful superordinate-basic phase 2 condition and the artificial superordinate-basic phase 2 condition showed a significant main effect of training block (F (3, 243) = 20.33, p < .001) an interaction between group and training block ($\underline{F}$ (3, 243) = 10.25, $\underline{p}$ < .001) as well as a main effect of group ($\underline{F}$ (1, 81) = 20.20, $\underline{p}$ < .001). Participants in the artificial superordinate-basic phase 2 condition were faster processing standard items on their respective first ($\underline{t}$ (81) = 24.88, $\underline{p}$ < .001) second, ($\underline{t}$ (81) = 11.16, $\underline{p}$ < .001) third, ($\underline{t}$ (81) = 13.29, $\underline{p}$ < .001), and fourth ($\underline{t}$ (81) = 9.70, $\underline{p}$ < .003) training blocks.

In sum, the artificial group was faster than the meaningful when transferring from one taxonomic level to another regardless of type of structure first learned. Clearly instructions failed to boost knowledge effects associated with the meaningful label.

The next set of comparisons examine phase 1 and phase 2 differences between artificial and meaningful groups. These comparisons include superordinate and basic level phase 1 conditions, as well as superordinate and basic level phase 2 conditions.

First, focusing on the artificial superordinate-superordinate phase 1 condition with the meaningful superordinate-superordinate phase 1 condition results showed a significant main effect of training block ($\underline{F}$ (3, 117) = 36.36, $\underline{p}$ < .001) and an interaction between training block and group condition ($\underline{F}$ (3, 117) = 3.18, $\underline{p}$ < .03). The main effect of group was statistically non-significant ($\underline{F}$ (1, 39) = 1.98, $\underline{p}$ > .17). Participants in the artificial label group processed items faster on training block 1 ($\underline{t}$ (39) = 2.26, $\underline{p}$ < .03).

Second, when comparing the artificial basic-basic phase 1 condition with the meaningful basic-basic label phase 1 condition, results showed a significant main effect of training block ($\underline{F}$ (3, 117) = 26.66, $\underline{p}$ < .001) an interaction between group and training block ($\underline{F}$ (3, 117) = 9.88, $\underline{p}$ < .001) as well as a main effect of group ($\underline{F}$ (1, 39) = 34.16, $\underline{p}$ < .001). Participants in artificial basic-basic group were faster than participants in the meaningful basic-basic group when processing items on training blocks 1 ($\underline{t}$ (39) = 5.28, $\underline{p}$

< .001), 2 (t (39) = 4.23, p < .001), 3 (t (39) = 4.77, p < .001) and 4 (t (39) = 3.70, p < .001).

Third, when comparing the artificial superordinate-superordinate phase 2 condition with the meaningful superordinate-superordinate phase 2 condition results showed a significant main effect of training block (F (3, 117) = 3.26, p < .03). The interaction between training block and group condition was statistically non-significant, (F (3, 117) = .43 p > .73) as was the main effect of group condition (F (1, 39) = .01, p > .93).

Finally, focusing on the artificial basic-basic phase 2 condition and the meaningful basic-basic label phase 2 condition results showed a significant main effect of training block (F (3, 117) = 11.40, p < .001). The interaction between training block and group condition was statistically non-significant (F (3, 117) = .96, p > .41) as was the main effect of group condition (F (1, 39) = 1.46, p > .24).

In sum, the artificial label group was reliably faster processing items in superordinate phase 1 and basic phase 1 conditions. This finding is surprising given the expectation of prior knowledge effects for the meaningful label.

In Experiment 3, meaningful labels and instructions for feature relations failed to facilitate taxonomic transfer. Indeed, reaction times would suggest a negative affect on performance. Participants in the meaningful label condition were slower to categorize items than in the previous experiment. Secondly, no differences in correct categorizations were found between artificial and meaningful label groups. One finding of interest is the very clear difference in reaction time performance between groups. Specifically, the artificial label group was considerably faster than the meaningful label group, particularly when categorizing items for basic level categories. Additionally, it would seem that being given feature relations alone did not improve performance as findings for the artificial group here appear similar to those in Experiment 1. So neither labels, nor relations, nor their combination seem to help learning much.

Before exploring these results it is critical to note that accuracy ratings for basic level training are quite low. Participants in this experiment took two hours to complete categorizations for the basic-basic level group, and only achieved between 60 and 70 percent correct by task end. Some might argue that this finding renders interpretation of reaction times meaningless. However, it is important to consider the following. Concurrent four category comparisons are extremely difficult, but performance is still far above chance (25%). Furthermore, low accuracy rates do not negate the comparisons of interest in this experiment, specifically reaction times differences between artificial and meaningful label groups. Indeed, accuracy rates on some level validate current comparisons as no accuracy differences were found between the groups, yet very real reaction time differences exist (thus no reaction time vs. accuracy trade offs). Finally, most errors are within superordinate errors. That is, participants are more likely to confuse *saxophone* with *flute* than with *drum* or *bell*. Together these observations suggest that participants are detecting similarities between categories and possibly contrasting basic level categories belonging to the same superordinate category. Moreover, this would suggest that participants are not often merely responding to items by guessing. Otherwise within and outside (e.g., saying an item is *drum* when it is a *flute*) errors would likely occur equally often. Nevertheless, it is difficult to rule out guessing as a contributing factor to group performance (e.g., participants' accuracy is low and confusing *flute* with *saxophone* does not rule out the possibility that they are uncertain).

When considering reaction time differences, one observation of note is that reaction time differences between superordinate and basic levels differ depending on group. Differences between taxonomic levels are much smaller for the artificial than for the meaningful label group. Moreover, in comparison to the artificial label, meaningful label categorizations are substantially slower for the basic level and marginally slower for the superordinate level. These findings are of interest because they suggest that groups approach categorization differently.

Factors underlying these differences are likely several including prior expectations, label abstractness, dimensional qualities, and item structure. When considering prior expectations, one very real difference between groups is that one has prior knowledge of the categories whereas the other one does not. Taxonomic labels for the artificial group are without meaningful content, thus it is unlikely that participants distinguish differences between levels based on prior experiences attached to the category label. The basic level label SSX and superordinate level label AAX provide no clues as to the identity of the object. In contrast, people learn about musical instruments and their taxonomic relationship (e.g., see Osborne & Calhoun, 1998) at a very early age. Consequently, for meaningful label participants, a strong relationship exists between the category label and prior experience. They have a clear idea of what flutes and drums are and the kinds of features attached to these categories. Mention the category *flute*, and many will consider an instrument light in weight, silver in color, that was either played in the school orchestra or seen at the local symphony. Importantly, people may at times have difficulty translating abstract features such as resonating frequency into things they know a little about like pitch. Moreover, the occasional values of current items (e.g., 50kg. Flute) may at times be surprising. This may explain why meaningful label participants were slower categorizing basic level items. People have very clear expectations as to what attributes constitute known categories. When expectations are strongly held, people may find it difficult to readjust their expectations to think of categories in new ways (see e.g., Keil, 2003, for a similar but slightly different observation).

Interestingly, the strength of expectations about the attributes of the category may vary depending on taxonomic level. Support for this idea follows from greater reaction time differences between taxonomic levels for the meaningful group. Markman (1985) notes that superordinate categories are often treated as mass nouns rather than as count nouns. Count nouns are categories, like chair, flute, piccolo, which can be pluralized, counted and treated as individual objects. In contrast, superordinate categories are often treated as homogenous masses, which cannot be counted, and are used to refer to collections of multiple items. The tendency to treat superordinate categories as homogenous masses may be even greater when attributes are abstract as in this experiment, because both the label and the attribute fail to clearly identify the category. As a result, participants may decide the category is difficult to know with certainty and expend less effort (in comparison to when categorizing a the basic level) deciding item membership. That is, they are more likely to make classification decisions based on how item information generally fits the category.

In contrast, when considering basic level count nouns the relationship between the attribute and the category is relatively clear. The category *saxophone* denotes specific attributes (e.g., *reeds, brass* etc.). However, in this experiment the mapping between the attribute and basic level category is not so clear; participants have a clear idea of the category identity but the attributes are abstract. Thus, participants know enough about the category to attempt categorization, however as attributes are abstract and inconsistent with prior experiences, additional time and effort is required to decide item membership. Moreover, when categories are familiar people tend to believe they know more about the category than actually do (see e.g., Kiel, 2003). Thus in the basic level condition, participants may be more challenged to resolve the relationship between the abstract attribute and the category.

One final factor contributing to reaction time differences between taxonomic levels is item structure. Dimensional values for items vary depending on taxonomic level with a greater number of dimensional values belonging to superordinate categories than basic categories. Moreover, a greater number of dimensional values repeat themselves over items for superordinate than basic levels. Thus, participants may simply find it easier to categorize items for superordinate categories.

In sum, the main finding in this experiment is that meaningful label participants treat some taxonomic levels differently from artificial label participants. When categorizing at either taxonomic level artificial label participants appeared to categorize without much consideration for how items relate to the category label. In contrast, at least when considering the basic level, prior experiences seem to heavily influence categorizations by the meaningful label group. Having specific prior expectations for basic level categories, meaningful participants appeared surprised by abstract features and took longer to categorize items. Expectations for superordinate categories are likely general and abstract attributes used in this experiment do not necessarily have to attach to any one category. As a result participants may have given less consideration as to the identity of the category.

In the next experiment, additional items are introduced allowing further examination of how prior expectations and item structure affect performance. Instructions presented in the current experiment will not be included in the next, as they failed to boost knowledge associated with the meaningful labels.

Chapter V

Experiment 4

Surprisingly, the results of the previous experiments revealed no learning advantage for meaningful label groups. Indeed, reaction time performance was poorer for the meaningful label group than for the artificial. Moreover, reaction time differences between taxonomic levels were greater for the meaningful group. One possibility is that these results are an artifact of instructions introduced in Experiment 3 negatively affecting performance for the meaningful group. That is, perhaps the instructions forced participants to look deeper into relations between attributes, thus resulting in extended examination, but not allowing participants to clearly see feature connections. What is uncertain is whether all or only some of these factors influences categorization.

Research (see e.g., Kaplan & Murphy, 2000; Murphy & Kaplan, 2000) generally shows that facilitation affects of prior knowledge increase as number of features related to the category increase. Kaplan and Murphy (2000) compared category learning with mixed theme features to category learning with intact theme features. In the intact theme condition participants learned about pairs of categories with features that were consistent with prior experience. For example, the features for one category were related to arctic vehicles and the features of the other category were related to jungle vehicles. In the mixed theme condition the categories were mixed so that a category might have 50% arctic features and 50% jungle features. Kaplan and Murphy reported that learning was worse in the mixed theme condition, suggesting there was more facilitation due to prior knowledge in the intact theme condition.

The present research does not use thematic features, however the idea that facilitation affects of prior knowledge increase, as the number of category features consistent to increase is relevant. In this experiment, two additional items are introduced. These items, along with standard items used in previous experiments, differ in number of features related to the category. The general idea is that items having more category consistent features will benefit more from prior knowledge than those items having fewer features. The other advantage is (failing facilitation affects of prior knowledge) these items will provide a means of determining what factors are interfering with mapping of attributes to their category (see below).

In order to address these concerns two additional items are introduced in this experiment. Prototype and fifty-fifty items have the same dimensions as standard items. What differentiates these items is their structure. Regardless of taxonomic level or category, all dimensional values for prototypes belong to their category. For example, dimensional values are coded *111111* (see tables 4 and 6) for both the basic level category *flute*, and superordinate level category *wind instrument*. Thus, all dimensional values for this item (and other prototype items) belong to their category, and none belong to a contrasting category. Fifty-fifty items (see table 5 and 7) are unusual in that when categorizing these items at the basic level they split their attributes half and half within each superordinate. For example, when considering “wind instrument” instantiations, dimensional values are coded *121212*. This item can best be described as half flute and half saxophone. Finally, most of dimensional values for standard items fall in their category. For example, dimensional values for the basic level category *flute* tend to have the value 1 on most dimensions. One instance of coding for this item is *111124*, showing that 4 dimensional values belong to category flute, and 2 belongs to contrasting categories. When categorizing this item for the superordinate category *wind instrument* 5 dimensions belong to the category *wind instrument* and 1 belong to the contrasting category.

Based on prior research (e.g., Kaplan & Murphy, 2000), the following is expected for item comparisons. First focusing on the superordinate meaningful condition, prototypes, having the greatest number of features consistent with their category, are expected to outperform standard and fifty-fifty items. Poorest performance is expected for fifty-fifty items as their structure fails to belong to any one category. With one qualification, similar predictions are expected for the artificial group. As there are clear structural differences between prototype and standard items, performance for these items should follow a similar pattern to that of the meaningful group. Performance for fifty-fifty items may follow a different pattern. When categorizing these items into superordinate categories all features belong to their category. Unlike participants in the meaningful group who have clear expectations for items, participants in the artificial group most probably have none, and thus may categorize solely based on number of features correctly predicting the category. In which case, performance for fifty-fifty may not differ all that much from prototypes.

The addition of prototype and fifty-fifty items are also important for examining groups differences discussed in Experiment 3. For example, as noted above all dimensional values for prototype items transfer correctly to their category regardless of taxonomic level. Moreover, dimensional values are consistent with what one would expect given their category (a *flute* weighs what one would expect it to weigh). Thus, inconsistencies associated with dimensional values are no longer a factor affecting performance for these items. While the influence of dimensional values has been removed, the abstract nature of the dimensions still remains. If abstract dimensions were a factor affecting group differences in the previous experiment, then mapping of prototype information to categories will remain difficult for the meaningful label and the usual facilitation affects associated with prior knowledge will fail to emerge. Similarly, given a negative influence of dimensions, differences between taxonomic levels should be greater for the meaningful group than for the artificial group when categorizing prototype items. This is because dimensional values for prototype items are the same when classified at either taxonomic level and thus should not impact performance. Alternatively, if dimensional values (e.g., an unusually high weight for a *flute*) were solely responsible for failing to find an affect of prior knowledge in Experiment 3, then meaningful group performance for prototype items should improve, as the prototype structure is now consistent with prior experiences. Finally, when considering fifty-fifty items the combination of dimensional values for these items are odd and inconsistent with prior expectations. Thus the artificial group is expected to outperform the meaningful group when classifying fifty-fifty items.

In sum, two additional kinds of items are introduced in this experiment. The inclusion of these items allows further exploration of issues addressed in earlier Experiments. One focus of prior experiments was on factors that boost prior knowledge. In Experiment 2, meaningful labels were introduced and failed to boost prior knowledge. Similarly, in Experiment 3, instructions describing feature relations also failed to facilitate prior knowledge. In this experiment, one expectation given an affect of prior experience is enhanced performance for the meaningful group when categorizing prototype items, and poorer performance when categorizing fifty-fifty items. Another reason for including these items is that given a failure to find an effect of prior knowledge, they allow for a closer examination of whether abstract dimensions, incongruent features, or some combination of the two affect how the meaningful group makes classification decisions.

Finally, prior research (e.g., Kaplan & Murphy, 1999) using thematic attributes has demonstrated that facilitation affects of prior knowledge increase as number of features related to the category increases. Similar to previous experiments, some items in the present experiment have greater numbers of features related to their category. This experiment differs in that item attributes are abstract. One question following from these observations is whether prior knowledge affects learning in a similar way when feature-based information is abstract.

Table 4

Abstract structures of prototype items for the basic level and superordinate level categories used in Experiments 4 and 5

Exemplar	D1	D2	D3	D4	D5	D6	Exemplar	D1	D2	D3	D4	D5	D6
	Category AAX/DAX							Category SSX/DAX					
P1	1	1	1	1	1	1	P2	2	2	2	2	2	2
	Category KKX/JAX							Category LLX/JAX					
P3	3	3	3	3	3	3	P4	4	4	4	4	4	4

Note. *Each prototype for basic level (P1-P4) and superordinate level categories (P1-P4)* *are represented by a row in the table*

Table 5

Abstract structures of fifty-fifty items for the basic level and superordinate level categories used in Experiment 4 and 5

Exemplar	D1	D2	D3	D4	D5	D6	Exemplar	D1	D2	D3	D4	D5	D6
	Category AAX/DAX							Category SSX/DAX					
F1	1	2	1	2	1	2	F2	2	1	2	1	2	1
	Category KKX/JAX							Category LLX/JAX					
F3	3	4	3	4	3	4	F4	4	3	4	3	4	3

Note. *Each fifty-fifty item for basic level (F1-F4) and superordinate level categories (F1-F4) are represented by a row in the table*

Table 6

Prototypes items presented to meaningful and artificial label groups

Dimension	*Prototypes used for Basic and Superordinate Levels*	
	FLUTE/AAX	SAX/SSX
1	weight = .4kg	weight = 1kg
2	complexity = 5p	complexity = 6p
3	internal volume = 24cu	internal volume = 76cu
4	energy required = 107e	energy required = 130e
5	resonant frequency = 180db	resonant frequency = 162db
6	total number of possible objects = 8ob	total number of possible objects = 10ob

Dimension	DRUM/KKX	BELL/LLX
1	weight = 42kg	weight = 55kg
2	complexity = 3p	complexity = 2p
3	internal volume = 821cu	internal volume = 1009cu
4	energy required = 248e	energy required = 195e
5	resonant frequency = 73db	resonant frequency = 90db
6	total number of possible objects = 15ob	total number of possible objects = 18ob

Note. *These 4 prototypes are seen at both the basic and the superordinate category level*

Table 7

Fifty-Fifty items presented to artificial and meaningful label groups

Dimension	*Fifty-Fifty Items used for Basic and Superordinate Levels*	
	FLUTE/SAX	SAX/FLUTE
1	weight = .4kg	weight = 1kg
2	complexity = 6p	complexity = 5p
3	internal volume = 76cu	internal volume = 24cu
4	energy required = 107e	energy required = 130e
5	resonant frequency = 162db	resonant frequency = 180db
6	total number of possible objects = 8ob	total number of possible objects = 10ob
	DRUM/BELL	BELL/DRUM
1	weight = 42kg	weight = 55kg
2	complexity = 2p	complexity = 3p
3	internal volume = 821cu	internal volume = 1009cu
4	energy required = 195e	energy required = 248e
5	resonant frequency = 73db	resonant frequency = 90db
6	total number of possible objects = 18ob	total number of possible objects = 15ob

Note. *For Experiment 4 only dimensional values 2, 4, and 6 are switched, However,*

these items are counterbalanced for Experiment 5, so that group 1 views fifty-fifty items *in which dimensional values 2, 4, and 6 are switched, and group 2 views dimensional values in which 1, 3, and 5 are switched. Also note that when categorizing fifty-fifty items, participants view categories as Flute, Saxophone, Drum, and Bell. These same items are presented to the artificial label group (but with artificial labels)*

5.1 Method

5.1.1 Participants

Two hundred twenty six university undergraduates volunteered to participate in this experiment for partial course credit. In total 18 participants were removed from analyses for failing to meet learning criterion. Eleven participants were removed for performing below chance, with the remaining 7 were removed for exceeding reaction times of 3 seconds.

5.1.2 Materials and Design

In this experiment a ninth training block is introduced, which in addition to the 48 standard items included in the previous experiments, includes 4 prototype and 4 fifty-fifty items, one for each basic level category. All other materials are identical to that of previous experiments.

5.1.3 Procedure

Experiment 4 procedures replicate those of Experiment 1 and 2.

5.2 Results and Discussion

Note that as fifty-fifty items have no one-to-one correlation with basic level categories, results for these items are explored only for superordinate categories.

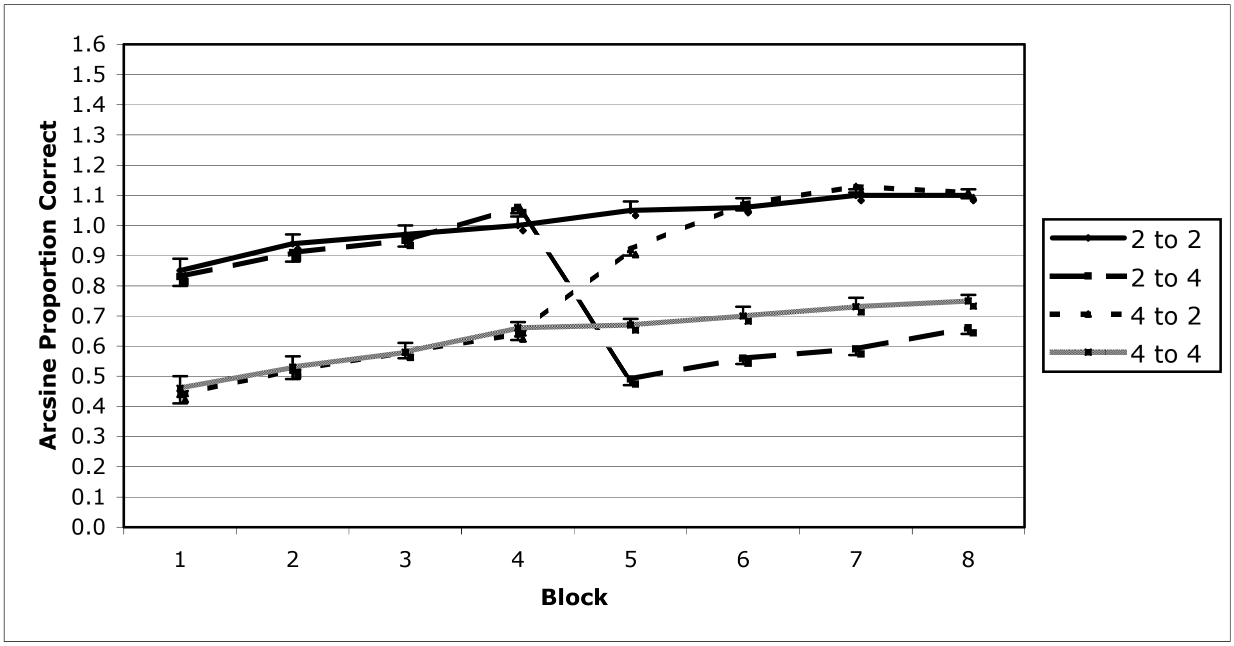
The key results for the first 8 training blocks are presented in Figures 9, 10, 11, and 12 separately for artificial and meaningful groups. As can be seen taxonomic affects are the same as the first three experiments, namely no taxonomic affects on accuracy. However, Phase 2 learning following exposure to basic level categories in phase 1 are overall faster than Phase 2 learning following exposure to superordinate categories in phase 1.

First focusing on accuracy ratings (see Figures 9 and 10) no clear differences are apparent between artificial and meaningful groups. Next focusing on reaction times (see Figures 11 and 12), performance is considerably slower for the meaningful label. Moreover, differences between basic and superordinate levels are much greater for the meaningful label. These findings replicate those of Experiment 3, and rule out the argument that findings in that study are solely attributable to an artifact of the instructions.

Next turning to performance for block 9 (see Figures 13, 14, 15, and 16), and focusing first on accuracy ratings (see Figures 13 and 14), prototypes (having congruent dimensional values) were favored over standard items, as expected. Interestingly, no differences are observed between prototype and fifty-fifty items. Moreover, this observation holds when focusing on reaction times (see Figures 15 and 16). This finding is inconsistent with the idea that meaningful label participants treat fifty-fifty items as either half flute/half saxophone or half drum/half bell. Finally, whether considering accuracy ratings or reaction times differences between groups appear slight. This observation is discussed in more detail in the discussion of this experiment.

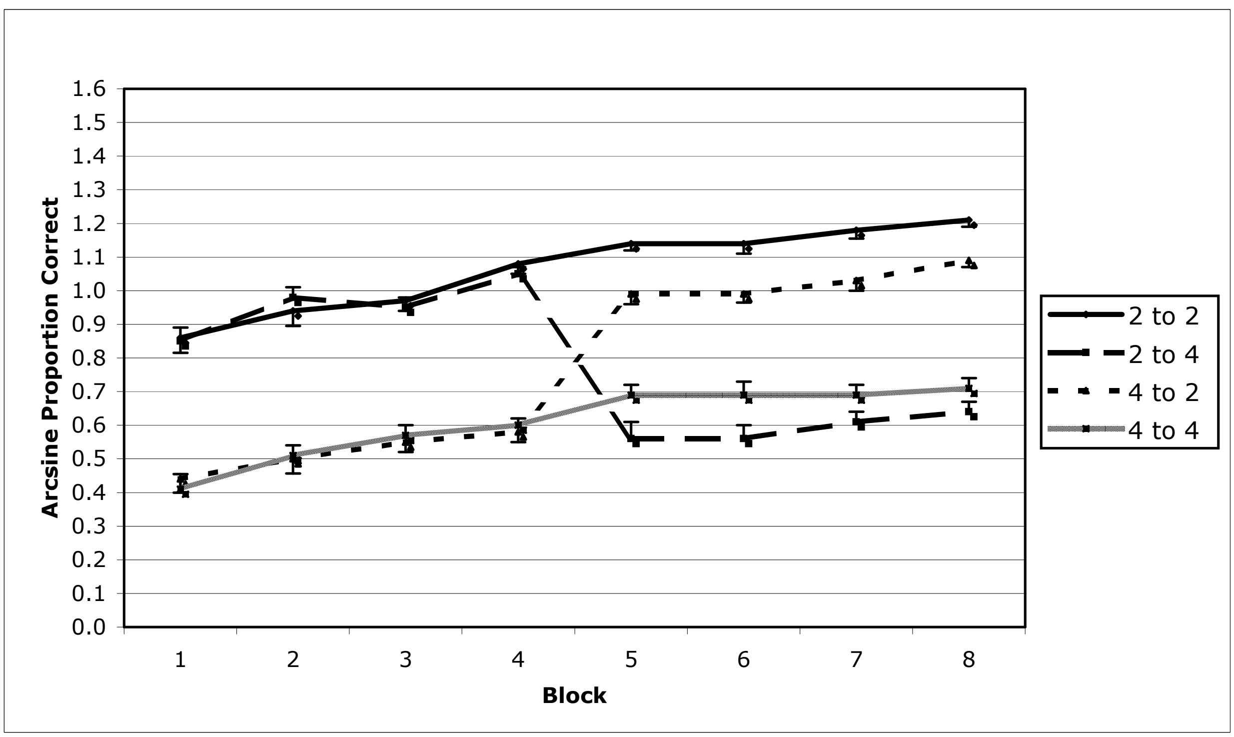
Analyses begin by first examining group differences and is followed up by independent group performance for block 9.

Figure 9.



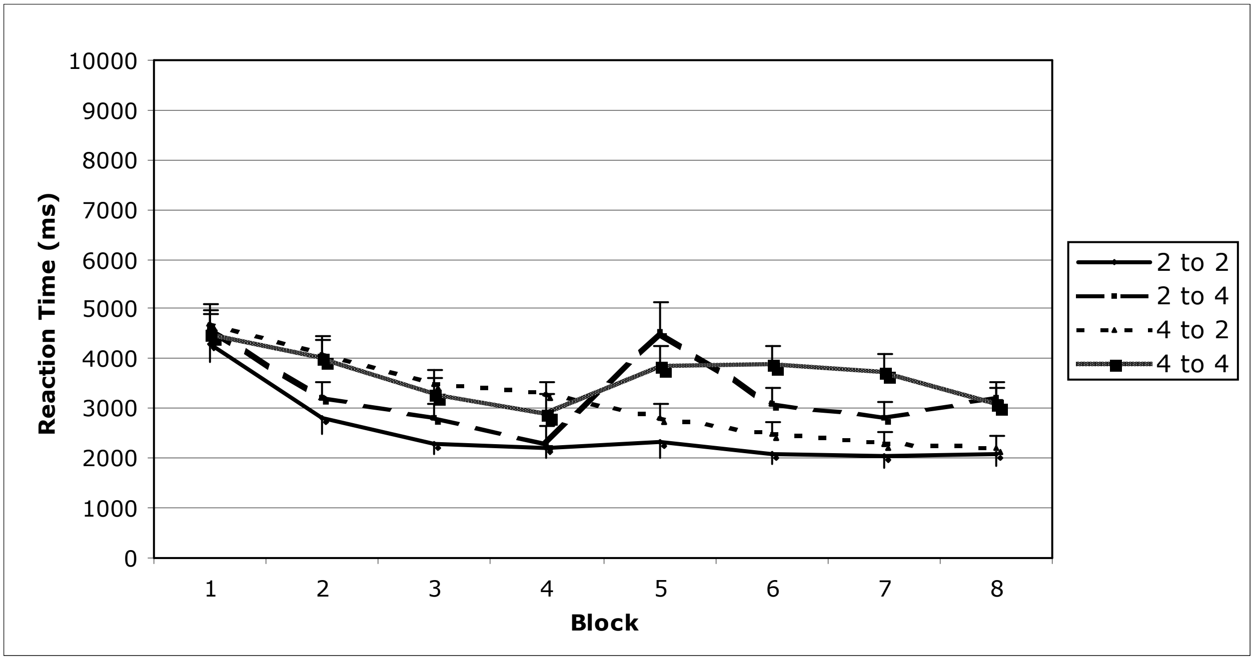
Learning of standard items for the artificial group in Experiment 4

Figure 10.



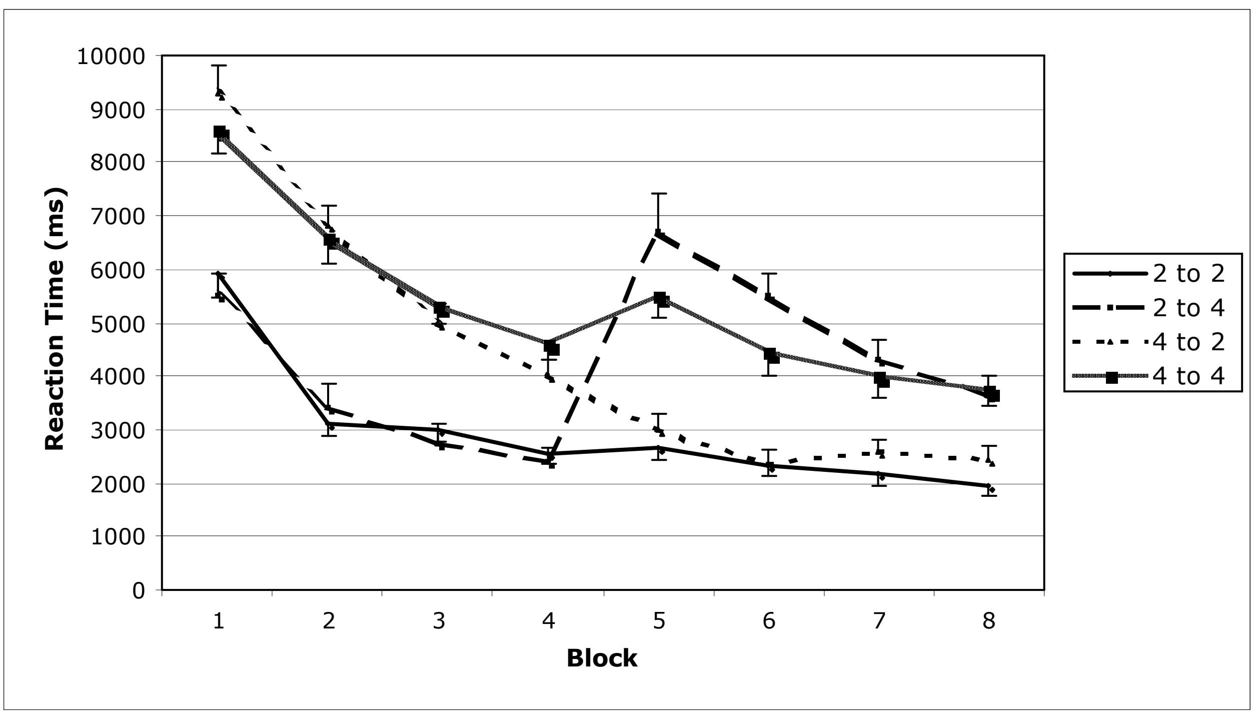
Learning of standard items for the meaningful group in Experiment 4

Figure 11.



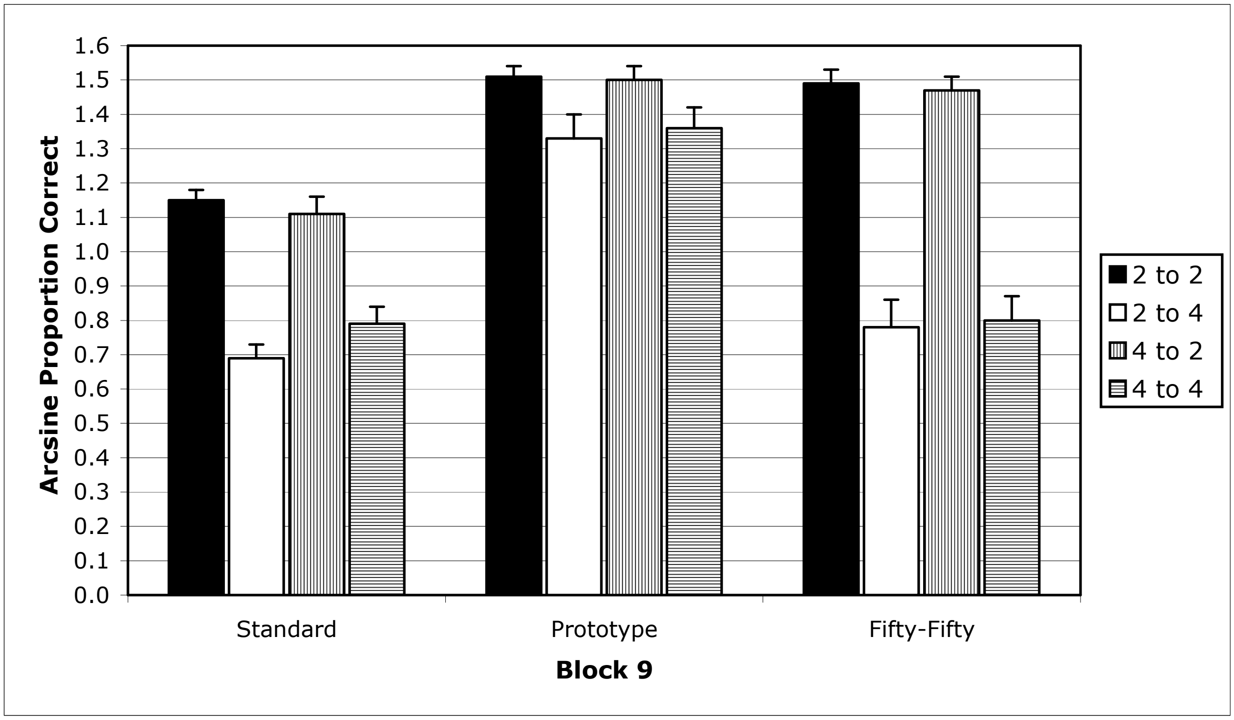
Standard item response times for the artificial group in Experiment 4

Figure 12.



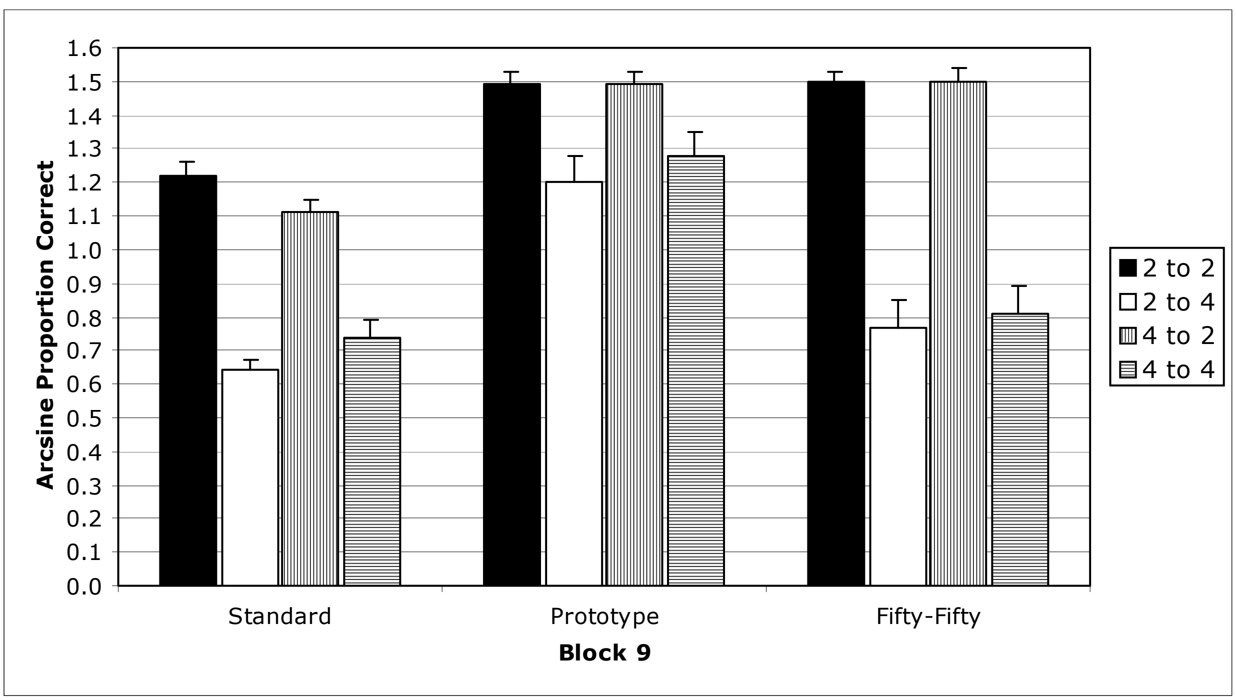
Standard item response times for the meaningful group in Experiment 4

Figure 13.



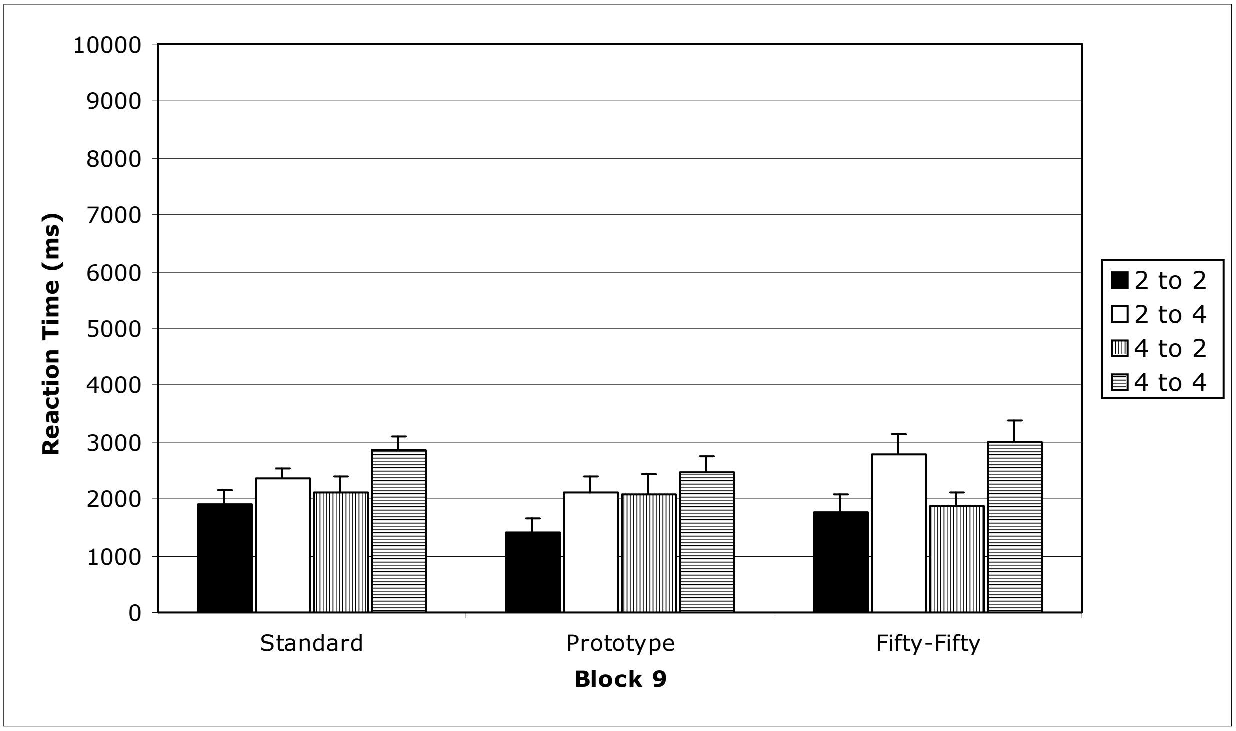
Item learning for the artificial group in Experiment 4

Figure 14.



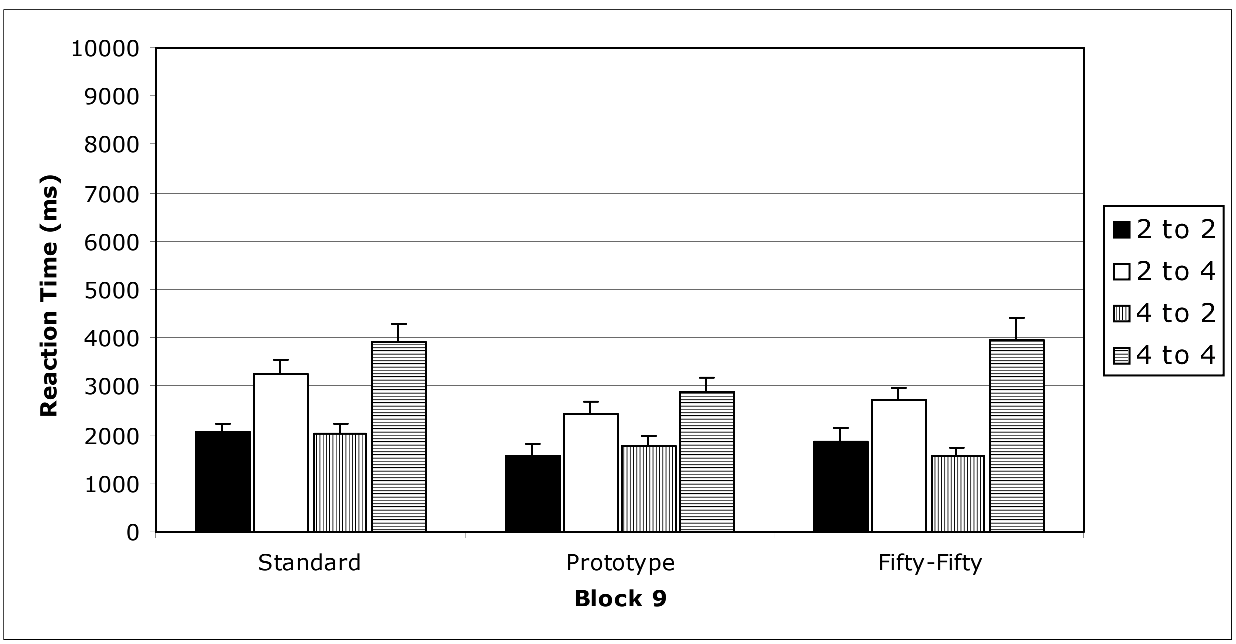
Item learning for the meaningful group in Experiment 4

Figure 15.



Item response times for the artificial group in Experiment 4

Figure 16.



Item response times for the meaningful group in Experiment 4

5.2.1 Artificial and meaningful group comparisons

The following sets of comparisons are important for verifying results found in the Experiment 3. In that Experiment no differences were found between artificial and meaningful groups on measures of accuracy. However, reaction time performance was overall much faster for the artificial label group. One factor that may have contributed to weak performance on the part of the meaningful group, particularly with respect to reaction time performance was the inclusion of instructions in that Experiment. Participants in the meaningful group having knowledge of categories may have spent more time trying to figure out how one feature was related to the other. Replicating the analysis performed in Experiment 3 is important for ruling out this possibility. The first sets of comparisons focus on learning of standard items over the first eight blocks of training. Block 9 comparisons for standard, prototype, and fifty-fifty items are explored later.

5.2.1.1 Accuracy

The first set of comparisons examine phase 1 and phase 2 differences between artificial and meaningful groups. These comparisons include superordinate and basic level phase 1 conditions, as well as superordinate and basic level phase 2 conditions. Findings favoring the meaningful group would show that meaning attached to the label boost classification for that group.

First, when comparing the artificial superordinate-superordinate phase 1 condition with the meaningful superordinate-superordinate phase 1 condition results showed a significant main effect of training block ($\underline{F}$ (3, 150) = 56.73, $\underline{p} < .001$). The interaction between group and training block was statistically non-significant (F (3, 150) = 1.95, p > .12) as was the main effect of group (F (1, 50) = 1.40, p > .24).

Next, when comparing the artificial basic-basic label phase 1 condition with the basic-basic meaningful label phase 1 condition results showed a significant main effect of training block (F (3, 150) = 111.24, p < .001). The interaction between training block and group was statistically non-significant (F (3, 150) = .90, p > .44) as was the main effect of group (F (1, 50) = 1.32, p > .26).

Third, when comparing the artificial superordinate-superordinate phase 2 condition with the meaningful superordinate-superordinate phase 2 conditions, results, showed a significant main effect of training block (F (3, 150) = 6.29, p < .02). The interaction between group and training block was statistically non-significant (F (3, 150) = .56, p > .65) as was the main effect of group (F (1, 50) = 2.24, p > .14).

Finally, focusing on the artificial basic-basic phase 2 condition and the meaningful basic-basic phase 2 condition, results showed a significant main effect of training block (F (3, 150) = 4.90, p < 003). The interaction between group condition and training block was statistically non-significant (F (3, 150) = .65, p > .58), as was the main effect of group condition (F (1, 50) = .21, p > .65).

In sum, analyses showed no differences between groups when learning superordinate and basic level categories. Thus, meaning attached to the meaningful label did not boost learning for that group. Moreover, these findings replicate those found in Experiment 3 and suggest that instructions presented in that experiment were not solely responsible for weak performance on part of the meaningful group.

5.2.1.2 Reaction Times

Possible outcomes for group reaction times are several. Faster performance on part of the meaningful group would not only suggest that meaning attached to the label boosted performance for that group, but that instructions presented in Experiment 3 adversely affected performance of the meaningful group. Conversely, a replication of findings in Experiment 3, that is slower reaction times for the meaningful label, would support the idea that meaning attached to label interacts adversely with abstract dimensions and incongruent dimensional values.

The next set of comparisons examine phase 1 and phase 2 differences between artificial and meaningful groups. These comparisons include superordinate and basic level phase 1 conditions, as well as superordinate and basic level phase 2 conditions.

First, when comparing the artificial superordinate-superordinate phase 1 condition with the meaningful superordinate-superordinate phase 1 condition, results showed a significant main effect of training block ($\underline{F}$ (3, 150) = 56.41, $\underline{p} < .001$). The interaction between training block and group condition was statistically non-significant ($\underline{F}$ (3, 150) = 3.23, $\underline{p} > .06$) as was the main effect of group condition ($\underline{F}$ (1, 150) = 1.18, $\underline{p} > .28$).

Second, focusing on the artificial basic-basic phase 1 condition and the meaningful basic-basic phase 1, results showed significant main effect of training block, $\underline{F}$ (3, 150) = 54.20, $\underline{p} < .001$, a significant interaction between training block and group condition ($\underline{F}$ (3, 150) = 9.50, $\underline{p} < .001$) as well as a main effect of group condition ($\underline{F}$ (1, 50) = 23.54, $\underline{p} < .001$). The artificial basic phase 1 condition was significantly faster processing items on blocks one ($\underline{t}$ (50) = 5.90, $\underline{p} < .001$), two ($\underline{t}$ (50) = 3.58, $\underline{p} < .001$), three ($\underline{t}$ (50) = 3.23, $\underline{p} < .002$, and four ($\underline{t}$ (50) = 2.56, $\underline{p} < .01$).

Third, focusing on the artificial superordinate-superordinate phase 1 condition with the meaningful superordinate-superordinate phase 1 condition showed a significant main effect of training block (F (3, 150) = 9.50, $p < .001$). The interaction between training block and group condition was statistically non-significant (F (3, 150) = 2.58, $p > .11$) as was the main effect of group condition (F (1, 50) = .28, $p > .60$).

Finally, when comparing the artificial basic-basic phase 2 condition with the meaningful basic-basic phase 2 condition results showed a significant main effect of training block (F (3, 150) = 9.10, $p < .001$). The interaction between training block and group condition was statistically non-significant (F (3, 350) = 2.44, $p > .07$), as was the main effect of group condition (F (1, 50) = 1.52, $p > .22$).

In sum, findings replicate those of Experiment 3. Participants in the meaningful group performed reliably slower than participants in the artificial group, but only when learning the basic-basic phase 1 condition. This finding suggests that meaning attached to the meaningful label interacts adversely with prior expectations of participants.

5.2.3 Artificial and meaningful label group comparisons for block 9.

The following analyses compare group differences when learning standard, prototype, and fifty-fifty items at block 9. As prototype items are congruent with participant's prior expectations one would expect enhanced performance on part of the meaningful group for these items. However, if abstract dimensions interact adversely with prior experiences, weaker performance (or equal performance) is expected on part of the meaningful group for these items.

5.2.3.1 Accuracy

First, when comparing the artificial superordinate-superordinate phase 2 condition with the meaningful superordinate-superordinate phase 2 condition, results showed a significant main effect of item type, (F (2, 100) = 57.87, $p < .001$). The interaction between group and item type was statistically non-significant (F (2, 100) = 1.34, $p > .27$), as was the main effect of group (F (1, 50) = 1.84, $p > .18$).

Second, focusing on the artificial basic-superordinate phase 2 condition and the meaningful basic-superordinate phase 2 condition, results, showed a significant main effect of item type (F (2, 100) = 49.35, $p < .001$). The interaction between group and item type was statistically non-significant (F (2, 100) = .10, $p > .91$) as was the main effect of group (F (1, 50) = .15, $p > .70$).

Third, when comparing the artificial basic-basic label phase 1 condition with the basic-basic meaningful label phase 1, results showed a significant main effect of item type (F (2, 100) = 50.76, $p < .001$). The interaction between training block and group was statistically non-significant (F (2, 100) = .28, $p > .76$) as was the main effect of group (F (1, 50) = .38, $p > .55$).

Finally, focusing on the artificial basic-basic phase 2 condition and the meaningful basic-basic phase 2 condition, results showed a significant main effect of training block (F (2, 100) = 33.80, $p < .001$). The interaction between group condition and training block was statistically non-significant (F (2, 100) = .56, $p > .58$) as was the main effect of group condition (F (1, 50) = .54, $p > .47$).

In sum, no differences were found between groups when learning prototype, fifty-fifty, and standard items. This would suggest that even when features are congruent with prior experience as with prototype items, the meaningful group fails to benefit.

5.2.3.2 Reaction times

First, focusing on the artificial superordinate-superordinate phase 2 condition and the meaningful superordinate-superordinate phase 2 condition results showed a significant main effect of item type, $\underline{F}$ (2, 100) = 2.86, $\underline{p} < .001$. The interaction between group condition and item type was statistically non-significant, $\underline{F}$ (2, 100) = .08, $\underline{p} > .93$, as was the main effect of group condition, $\underline{F}$ (1, 50) = .16, $\underline{p} > .69$.

Second, when comparing the artificial basic-superordinate phase 2 condition and the meaningful basic-superordinate phase 2 condition, results showed a significant main effect of item type ($\underline{F}$ (2, 100) = 17.42, $\underline{p} < .001$). The interaction between group condition and item type was statistically non-significant ($\underline{F}$ (2, 100) = 3.01, $\underline{p} > .06$) as was the main effect of group condition ($\underline{F}$ (1, 50) = 2.38, $\underline{p} > .13$).

Third, when comparing the basic-basic phase 2 condition with the meaningful basic-basic phase 2 condition, results showed a significant main effect of item type ($\underline{F}$ (2, 100) = 3.64, $\underline{p} < .03$). The interaction between group and item type was statistically non-significant ($\underline{F}$ (2, 100) = .58, $\underline{p} > .56$) as was the main effect of group condition ($\underline{F}$ (1, 50) = 1.83, $\underline{p} > .18$). Note, that an independent sample t-test showed that the artificial label group was faster processing standard items ($\underline{t}$ (50) = 2.16, $\underline{p} < .04$).

Finally, when focusing on the artificial superordinate-basic phase 2 condition the meaningful superordinate-basic phase 2 condition results determined that the main effect of item type was statistically non-significant ($\underline{F}$ (2, 100) = 2.86, $\underline{p} > .06$) as was the interaction between group condition and item type (F (2, 100) = 1.84, $p > .16$) and the main effect of group condition (F (1, 50) = 1.50, $p > .23$).

In sum, the only reaction time differences found favored the artificial group when learning standard items at the basic-basic phase 2 condition. Results for both accuracy and reaction times suggest that even when items are congruent with prior experience, as in the case of prototype items, the meaningful group performance fails to exceed the performance of the artificial group. It's possible that either abstract dimensions and/or incongruent dimensional values negatively impacted performance of the meaningful group. In the case of prototype items, the only item characteristics not controlled for were the abstract dimensions. Thus, it's possible that abstract dimensions interfered with or limited access to prior experiences. However, it is important to note that with respect to reaction times the meaningful group also performed poorly on standard items. This may indicate that dimensional values also contributed to meaningful group performance. The next section set of analyses separate influences of abstract dimensions and dimensional values on group performance.

5.2.4 Block 9 item comparisons for individual groups

5.2.4.1 Artificial label group

The following sets of analyses examine differences between items for block 9. These analyses are important for exploring how differences in item structure affect learning of items. Participants are expected to prefer the structural qualities of prototype items as compared to standard items. Expectations for prototype and fifty-fifty items are less clear. If participants categorized based on the number of dimensional values that belong to superordinate categories, then small differences are expected for these items.

On the other hand, if they classify based on prior knowledge poorer performance is expected for fifty-fifty items. Comparisons are made first for the artificial label and then for the meaningful label.

5.2.4.2 Accuracy

As can be seen in figure 14, prototype items were classified better than standard items in both the superordinate-superordinate phase 2 (t (25) = 6.94, $p < .001$) and basic-superordinate phase 2 conditions (t (25) = 7.14, $p < .001$). Moreover, prototype items were also classified better than standard items in both the basic-basic phase 2 (t (25) = 9.64, $p < .001$) and superordinate-basic phase 2 conditions (t (25) = 9.64, $p < .001$). However, no differences were found between prototype and fifty-fifty items in either the superordinate-superordinate phase 2 (t (25) = .57, $p = .57$) and the basic-superordinate phase 2 conditions (t (25) = .87, $p > .39$).

Finally, fifty-fifty items were classified better than standard items in both the superordinate-superordinate phase 2 (t (25) = 5.97, $p < .001$) and the basic-superordinate phase 2 conditions (t (25) = 6.79, $p < .001$).

When considering accuracy findings at block 9 performances, for prototype and fifty-fifty items were generally better than for standard items. Moreover, no differences were found between prototype and fifty-fifty items, even in the basic-superordinate phase 2 conditions where participants have just been taught categories that correspond to the prototype items but not to the fifty-fifty items. The findings are consistent with the idea that participants are making classification decisions based on the number of dimensional values that correctly predict the category.

5.2.4.3 Reaction times

As can be seen in figure 15, prototype items were classified faster than standard items when classifying at the superordinate-superordinate phase 2 level (t (25) = 2.47, $p <$ .02). However, no reaction time differences were found between these items when classifying in the basic-superordinate phase 2 condition (t (25) = .24, p = .82). Moreover, no reaction time differences were found between prototypes and standard items in either the basic-basic phase 2 condition (t (25) = 1.99, $p >$ .06) or the superordinate-basic phase 2 condition (t (25) = .86, $p >$ .39). A null finding was also found when comparing reaction time differences between prototype and fifty-fifty items at both the superordinate-superordinate phase 2 condition, (t (25) = 1.92, $p >$ .07), and the basic-superordinate phase 2 condition (t (25) = .68, $p >$ .50).

Finally, no reaction time differences were found between standard and fifty-fifty items in either the superordinate-superordinate phase 2 condition (t (25) = .73, $p >$ .47) or the basic-superordinate phase 2 condition (t (25) = 1.47, $p >$ .15).

In sum, prototype items were classified faster than standard items at the superordinate-superordinate phase 2 level, otherwise no other differences were found between items.

5.2.5 Meaningful label group

Meaningful group participants are expected to classify prototype items better than both standard items. Moreover, in comparison to standard items and prototype items poorer performance is expected for fifty-fifty items. This is because the combination of dimensional values for fifty-fifty is inconsistent with participant's prior expectations for instruments.

5.2.5.1 Accuracy

As can be seen in figure 14, Prototype items were classified better standard items in both the superordinate-superordinate phase 2 condition (t (25) = 4.64, p < .001) and basic-superordinate phase 2 condition (t (25) = 8.12, p < .001). Furthermore, prototype items were classified better than standard items in both the basic-basic phase condition (t (25) = 8.00, p < .001) and superordinate-basic phase 2 condition, (t (25) = 7.63, p < .001). However, no differences were found between prototype and fifty-fifty items when classifying these items in the superordinate-superordinate phase 2 condition (t (25) = 1.81, p > .08) or the basic-superordinate phase 2 condition (t (25) = .27, p > .79).

Finally, fifty-fifty items were classified better than standard items in both the superordinate-superordinate phase 2 condition (t (25) = 6.89, p < .001) and the basic-superordinate phase 2 condition (t (25) = 7.81, p < .001).

When considering accuracy findings at block 9, performances for prototype and fifty-fifty items was generally better than for standard items. Moreover, no differences were found between prototype and fifty-fifty items. The findings are consistent with the idea that participants are making classification decisions based on the number of dimensional values that correctly predict the category. Thus participants do not appear to be making decisions based on prior experiences with instruments (otherwise a half flute/half saxophone would seem odd in comparison to an instrument that is all flute or mostly flute).

5.2.5.2 Reaction times

As can be seen in figure 14, prototype items were categorized faster than standard items in the superordinate-superordinate phase 2 condition (t (25) = 4.00, p < .001). However, no differences were found between these items when categorizing in the basic-superordinate phase 2 condition (t (23) = 1.33, $p > .82$). Furthermore, no reaction time differences were found between prototype and fifty-fifty items when categorizing in either the superordinate-superordinate phase 2 condition (t (25) = 1.88, $p > .07$) or the superordinate-basic phase 2 condition (t (25) = 1.26, $p > .22$).

Finally, fifty-fifty items were categorized faster than standard items in the basic-superordinate phase 2 condition (t (23) = 2.60, $p < .02$). However no reaction time differences were found between these items in the superordinate-superordinate phase 2 condition (t (25) = 1.18, $p > .25$).

In sum, prototype items were categorized faster than standard for the basic level conditions and for the repeated superordinate condition. Fifty-fifty items were categorized faster than standard for the basic-superordinate phase 2 condition. Importantly, no reaction time differences were found between prototype and fifty-fifty items. Thus, the meaningful group did not appear to treat fifty-fifty items differently from prototype items.

5.2.6 Differences between taxonomic levels

Then next sets of analyses examine mean differences between superordinate and basic level categories for artificial and meaningful groups. These analyses are important for differentiating the influence of prior expectations, dimensions, and item structure, on categorization. For example, slower responses on part of the meaningful group for prototype items would suggest that abstract dimensions negatively impacted performance. This is because the primary factor affecting performance for prototype items are abstract dimensions (structure for prototype items was held constant between taxonomic levels). However, slower responses to standard items on part of the meaningful group would suggest that both abstract dimensions and incongruent dimensional values negatively affected meaningful group performance (these items have both abstract dimensions and incongruent dimensional values). Finally, slower between taxonomic level responses to standard than to prototype items on would suggest that incongruent dimensional values are the primary factor affecting performance. The first set of analyses compares mean differences between basic-basic and superordinate-superordinate groups. The second set of analyses examines group mean differences between basic-superordinate and superordinate-basic phase 2 conditions.

First, when comparing differences between basic-basic and superordinate-superordinate groups for standard items results showed mean differences were smaller for the artificial label on blocks, 1, 2, 3, and 5 ($p < .02$). Comparisons for block 9 failed to find mean reaction time differences between groups for standard (t (50) = 1.65, $p > .24$) prototype (t (50) = 1.23, $p > .37$). Moreover, no between taxonomic level differences were found when comparing prototype and standard items (t (50) = .89, $p > .56$).

Next, when comparing differences between basic-superordinate and superordinate-basic phase 2 condition for standard items no mean differences were found between groups on block, 5, 6, 7, and 8 ($p > .77$). Moreover comparisons for block 9 also failed to reveal mean reaction time differences between groups for standard (t (50) = 1.83, $p > .07$) and prototype items (t (50) = 1.70, $p > .09$). No between taxonomic level differences were found when comparing prototype and standard items (t (50) = .72, $p > .53$).

That mean differences between taxonomic levels were smaller for the artificial than for the meaningful group when processing standard items indicates that something about the items' attributes negatively affected performance for the meaningful group. However, as no performance differences were found between groups when comparing prototype items with standard items, and when comparing prototype items alone, it is difficult to differentiate the extent to which abstract dimensions and incongruent dimensional values contributed to performance.

One main finding in this experiment is that prior knowledge had little influence on correct categorization. Participants who learned standard, prototype, and fifty-fifty items in the presence of meaningful labels failed to outperform participants who categorized the items with artificial labels. Moreover, consistent with Experiment 3, the processing of standard items in the presence of the meaningful label was substantially slower for the basic level (through the first several blocks of training) and slightly slower for the superordinate level. Additionally, mean response time differences between taxonomic levels for standard items were substantially smaller for the artificial group. Thus, as found in Experiment 3, meaningful label participants treated taxonomic levels differently than artificial label participants.

It is important to note that previous reviews of prior knowledge effects on categorization (e.g., Kaplan & Murphy, 2000; Spalding & Murphy, 1996) used thematic relations. Failure to find prior knowledge effects found in this study may in part be an artifact of weak connections between the features and an overt theme. In the context of musical instruments, the dimensions for the current items may be largely unfamiliar and not easily transferred to prior experiences. That response time comparisons between taxonomic levels and groups failed to differ on prototype items would seem to suggest this. This is because the main factor affecting between taxonomic level performances for prototype items is abstract dimensions (dimensional values are held constant between taxonomic levels for prototype items). Thus, it is possible that participants found deciphering of abstract dimensions difficult and they interfered with participant's ability to classify prototype items. However, this explanation is not entirely satisfactory. A stronger argument in favor of an influence of abstract dimensions would be decidedly slower responses for the meaningful group when classifying prototype items. Moreover, at least during initial stages of training clear differences were found between groups when classifying standard items, thus it is also possible that incongruent dimension values contributed to performance of the meaningful group. It is important to note that because prototype items were introduced in Block 9 comparisons between prototype items and standard items learned in Block 1 is not possible. For reasons addressed in Experiment 5, introducing prototype items into earlier blocks may provide a clearer picture as to whether abstract dimensional values adversely affect meaningful group performance.

Dimensional values may also influence performance in other ways. Surprisingly, no meaningful differences were found between fifty-fifty and prototype items when categorizing at the superordinate level. These findings are clearly inconsistent with the position that participants treat fifty-fifty items as incongruent with prior experiences, even if they have been taught basic level categories consistent with the prototype items and consistent with the fifty-fifty items. One reason for these findings may follow from the fact that dimensional values for prototype and fifty-fifty items classify in a similar way (see Tables 6 and 7) in that all dimensional values for prototype and fifty-fifty items transfer to the same superordinate category and none belong to the contrasting category. Moreover, dimensional values between prototype and fifty-fifty items are very close (e.g., the dimensional value for flute on the complexity dimension is 5kg. for the flute prototype and 6kg. for the flute/sax fifty-fifty item). Thus, at least when classifying these items at the superordinate level, participants may treat these items as very similar. As discussed in the next Experiment the tendency to rely on similarity judgments may be even greater when item properties are abstract.

Chapter VI

Experiment 5

Recent research (Heit & Bott, 2000) has shown that the facilitation affect of prior knowledge varies in magnitude over the course of category learning. This research demonstrated that benefits of prior experience emerged only after participants have been exposed to enough information. Results found thus far in this study suggest a slightly different trend. Specifically, when item attributes are unfamiliar (i.e., having both weak connections to one another and to an overt theme) mapping of item information is difficult for meaningful label participants particularly during initial stages of training. As an example, group reaction time differences were always much greater for standard items during the first several blocks of training. One reason for this finding may follow from meaningful participants finding abstract item information more surprising during initial stages of learning. This explanation may in part also explain why no reaction time differences were found between artificial and meaningful groups for prototype items in Experiment 4. In that Experiment prototype items were introduced in block 9.

In this experiment, prototype and fifty-fifty items are introduced into blocks 1 and 5, as well as block 9. Consistent with findings for standard items in earlier experiments one prediction in this experiment is that meaningful participants find mapping of item information more difficult when prototype and fifty-fifty items are first encountered. Reaction time differences between taxonomic levels are also expected to be greater during initial stages of learning. A finding of greater differences between taxonomic levels for prototype items by the meaningful group would support the idea that abstract item dimensions interfere with mapping of item information to prior experiences.

Importantly, dimensional values that are incongruent with prior experiences may also affect group differences. Greater differences between taxonomic levels for standard items than for prototypes by the meaningful group would also suggest that dimensional values interfere with mapping of item information. As participants become more familiar with item characteristics, group differences are expected to diminish.

An alternative prediction is an affect of prior knowledge. It is possible that participants' performance had reached asymptote by block 9, and prior knowledge affects were unable to manifest. That is, participants had already learned all they possibly could, leaving little room for group differences to emerge. This observation has greater relevance for superordinate categories where participants learned all but a few items. Another possibility is that after two hours of categorizing items participants' motivation was low and categorization was not a priority. Given the correctness of one or both of these observations one might predict affects of prior knowledge during earlier stages of learning when participants are motivated and still learning. Under these circumstances outcomes in this experiment would follow predictions made in the introduction of Experiment 4. For example, the meaningful label group performance for prototype items would be better than that for fifty-fifty items. This prediction follows from participants in the meaningful group finding items that are half flute/half saxophone and half drum/half bell inconsistent with their prior experiences.

Other predictions relate to generalization or the idea that participants notice the superordinate category while learning basic level categories. One would expect that if generalization contributes to categorization then following transfer from basic to superordinate condition only minor differences would be found between prototype and fifty-fifty items. That is, if participants learn that the flute category is similar to the saxophone category then the fact that the fifty-fifty item is a blend of the two, and has all dimensional values transferring to superordinate category, may lead fifty-fifty items to transfer just as well as prototype items (which at the basic level have all dimensional values belonging to one of the two categories). Moreover, standard items having fewer dimensional values transferring to the basic categories of flute and saxophone, and more dimensional values transferring to contrasting categories may transfer more poorly to the superordinate category than prototypes and fifty-fifty items.

In sum, the main question asked in this Experiment centers around outcomes that might occur when prototype and fifty-fifty items are introduced into earlier blocks of training. For example, will facilitation affects of prior knowledge manifest during initial training or alternatively will participants find abstract stimuli surprising resulting in prior knowledge negatively affect performance?

6.1 Method

6.1.1 Participants

Two hundred-fifty university undergraduates volunteered to participate in this experiment for partial course credit. In total 26 participants were removed from analyses for failing to meet learning criterion. Fourteen participants were removed for performing below chance, with the remaining 12 removed for exceeding reaction times of 30 seconds.

6.1.2 Materials and Design

The only difference between this experiment and Experiment 4 is that in addition to presenting prototype and fifty-fifty items in block 9, prototype and fifty-fifty items were introduced into blocks 1 and 5. One prototype and one standard item were presented in each category for blocks 1 and 5. Prototype and fifty-fifty items presented in block 9 were identical to those of previous blocks, but instead presenting one of these items for each category two of each were presented. All other materials were identical to those of previous experiments.

6.1.3 Procedure

Experiment 5 procedures replicate those of previous Experiment 4.

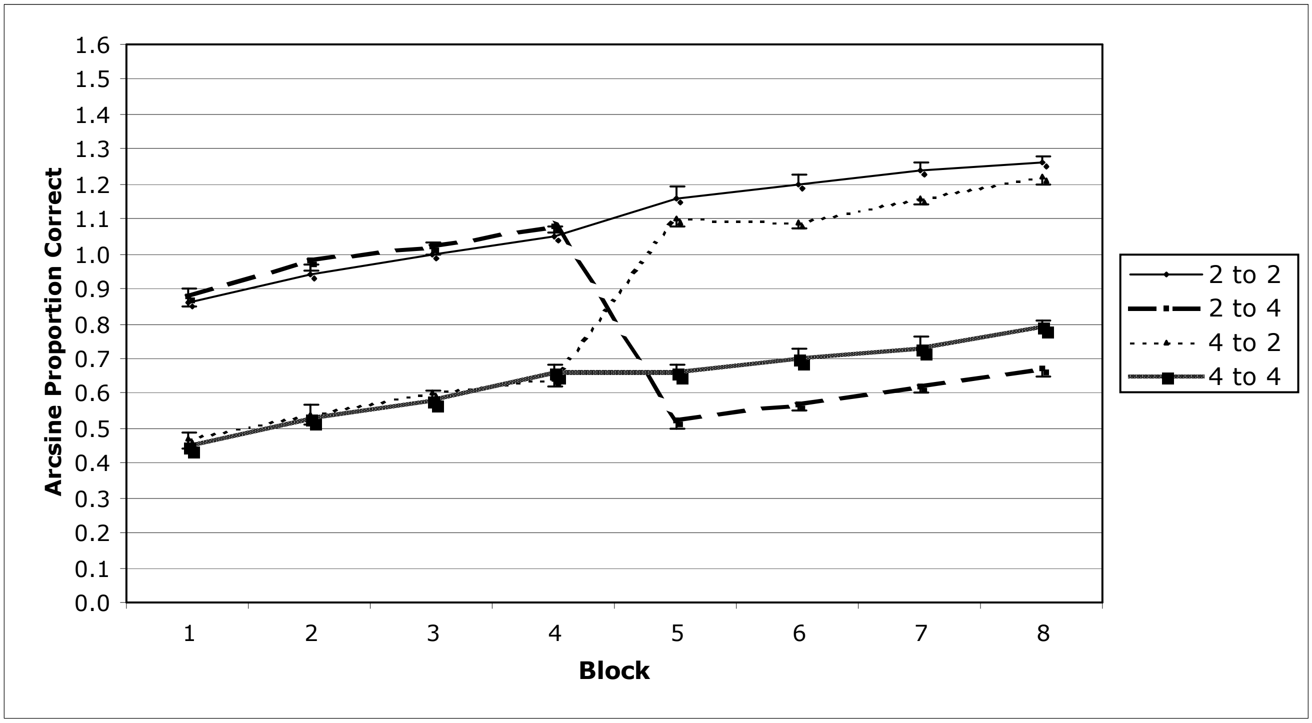
6.2 Results and Discussion

The key results are presented in Figures 17 - 32 separately for artificial and meaningful groups. Figures 17-20 depict results for standard items over the first eight blocks of learning. Replicating results found in previous experiments no observable differences are apparent between artificial and meaningful groups on measures of accuracy (see Figures 17 and 18). Moreover, as found in previous experiments, reaction times are generally faster for the artificial group when classifying at the basic level. Statistical results for these comparisons replicate those found in previous experiments and are therefore not presented in this experiment but are available upon request.

When focusing on accuracy ratings on blocks 1, 5, and 9 for all items (see Figures 21, 22, 25, 26, 29, 30), no clear differences are apparent between artificial and meaningful groups. This finding replicates those found in Experiment 4 and substantiates meaningful label participants having difficulty mapping unfamiliar item information. Next focusing on reaction times for blocks 1, 5, and 9 (see Figures 23, 24, 27, 28, 31, 32), performance is considerably slower for the meaningful label. Additionally, differences between meaningful basic and superordinate levels are greater for both standard items and prototype items. This observation further supports the idea that meaningful label participants have greater difficulty mapping unfamiliar information when first encountered.

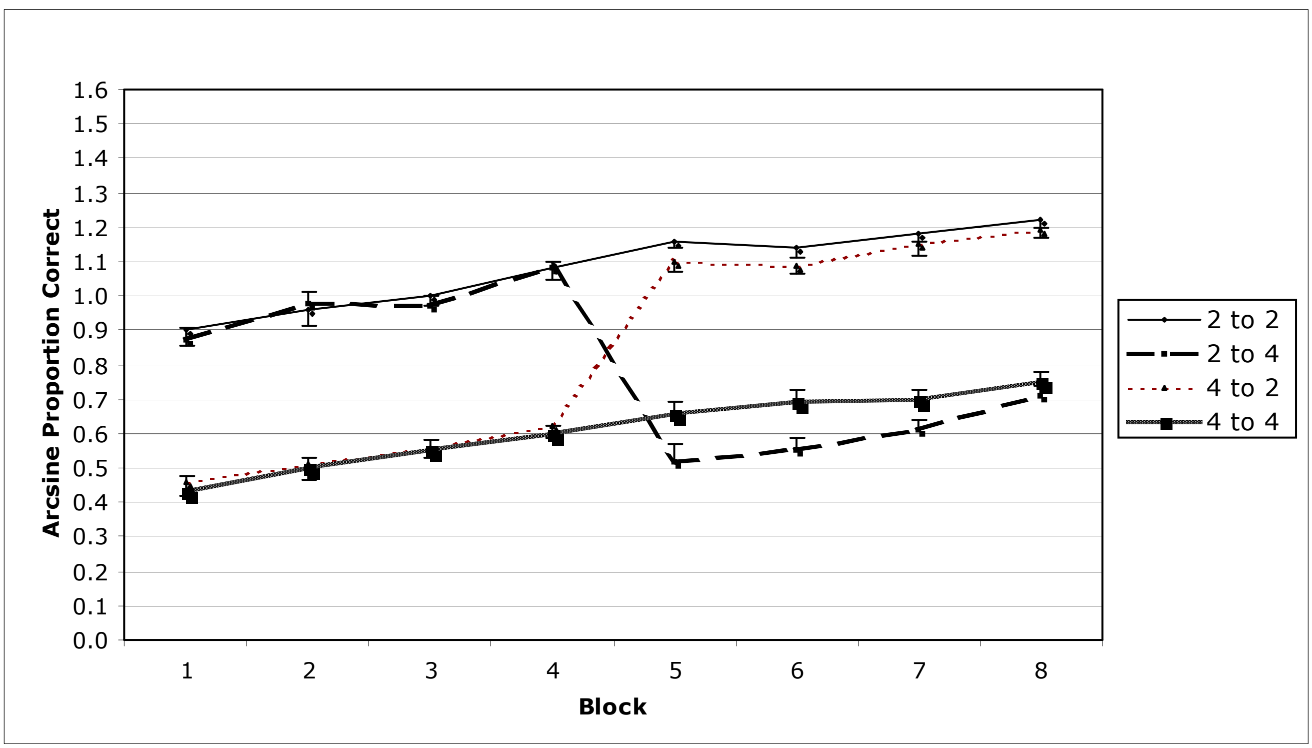
Finally, focusing on basic to superordinate level transfers, no observable differences are apparent between prototype and fifty-fifty items. This would suggest that participants are choosing to categorize these items largely based on how dimensional values generalize. That is, they appear to be learning at the basic level that dimensional values for prototype and fifty-fifty items belong to the same superordinate instantiations, and as a result are already inferring the superordinate categories.

Figure 17



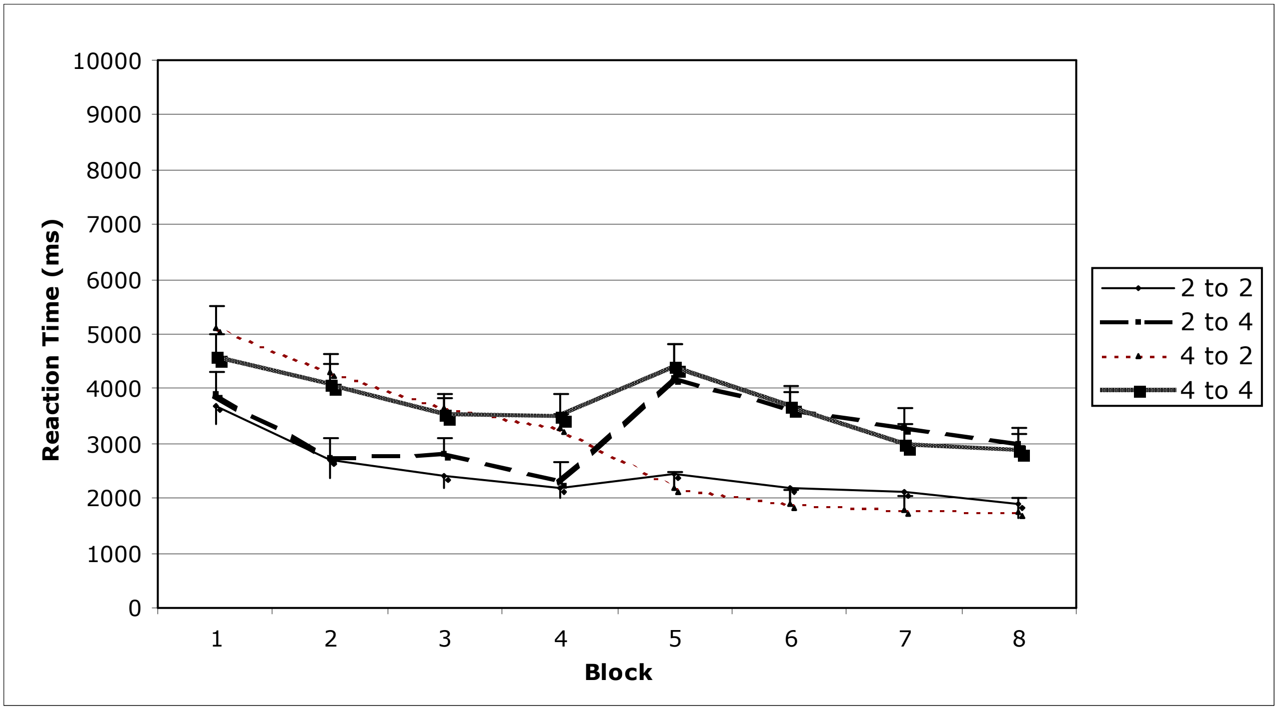
Learning of standard items for the artificial group in Experiment 5

Figure 18



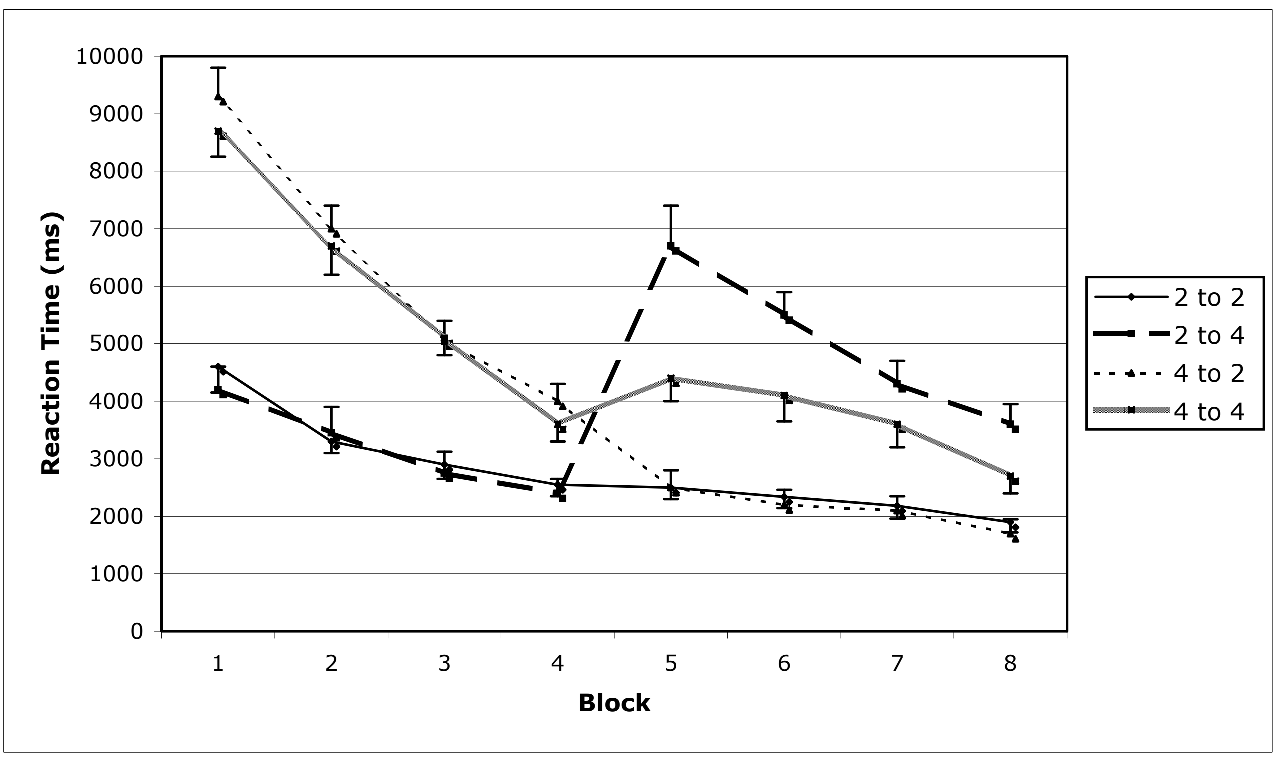
Learning of standard items for the meaningful group in Experiment 5

Figure 19



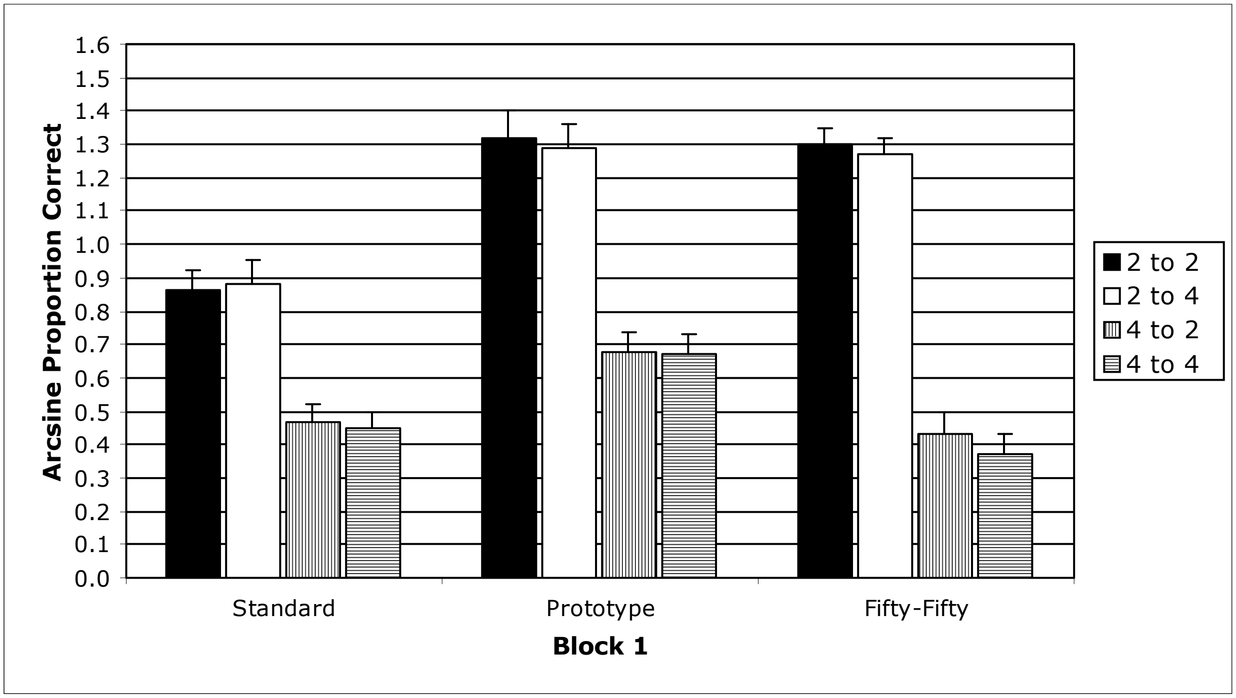
Standard item response times for the artificial group in Experiment 5

Figure 20



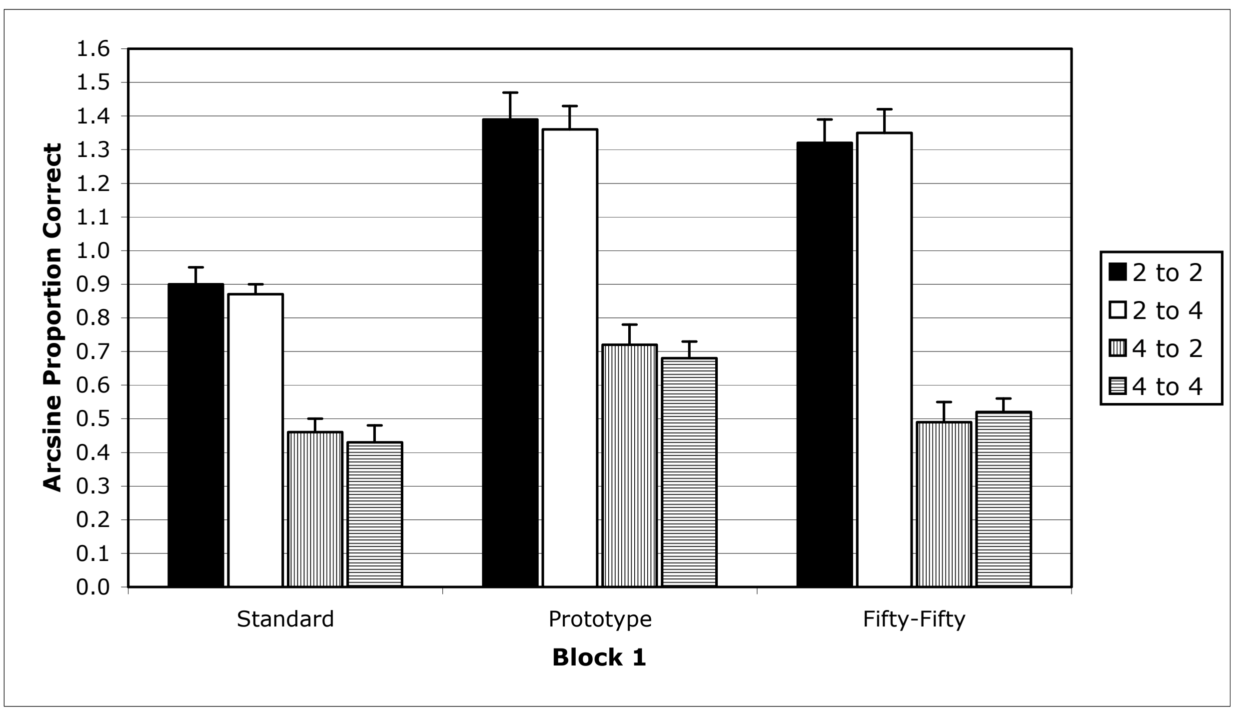
Standard item response times for the meaningful group Experiment 5

Figure 21.



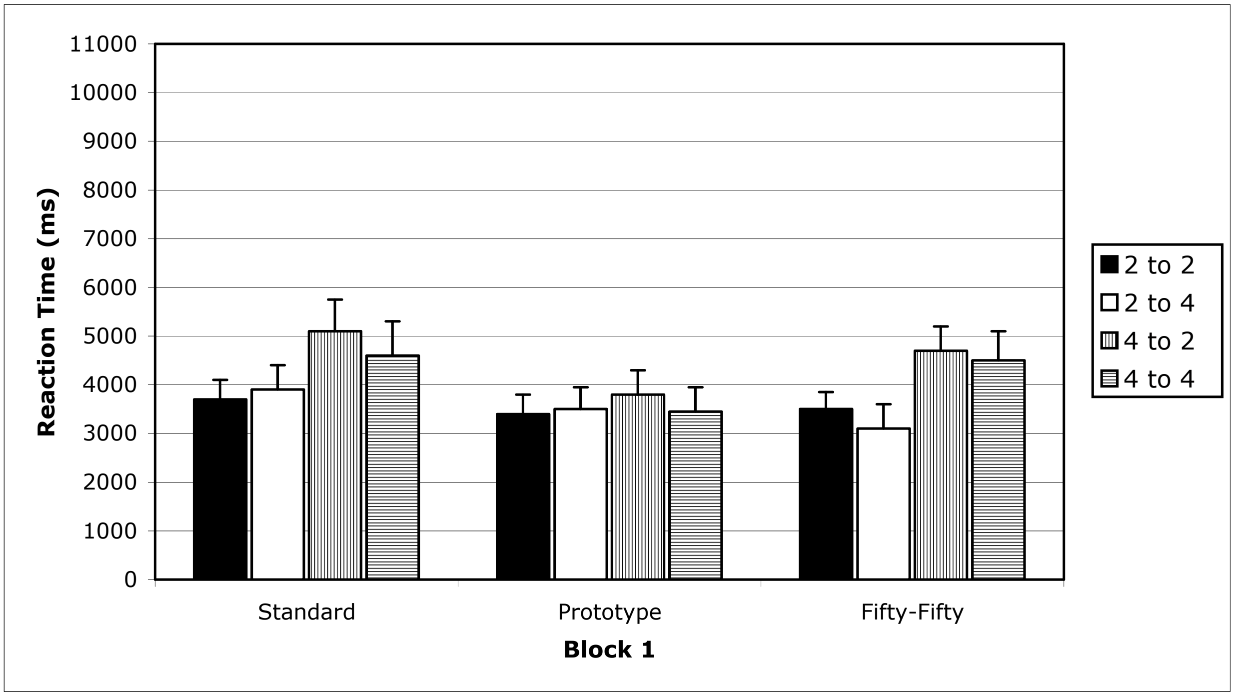
Item learning for the artificial group in Experiment 5

Figure 22.



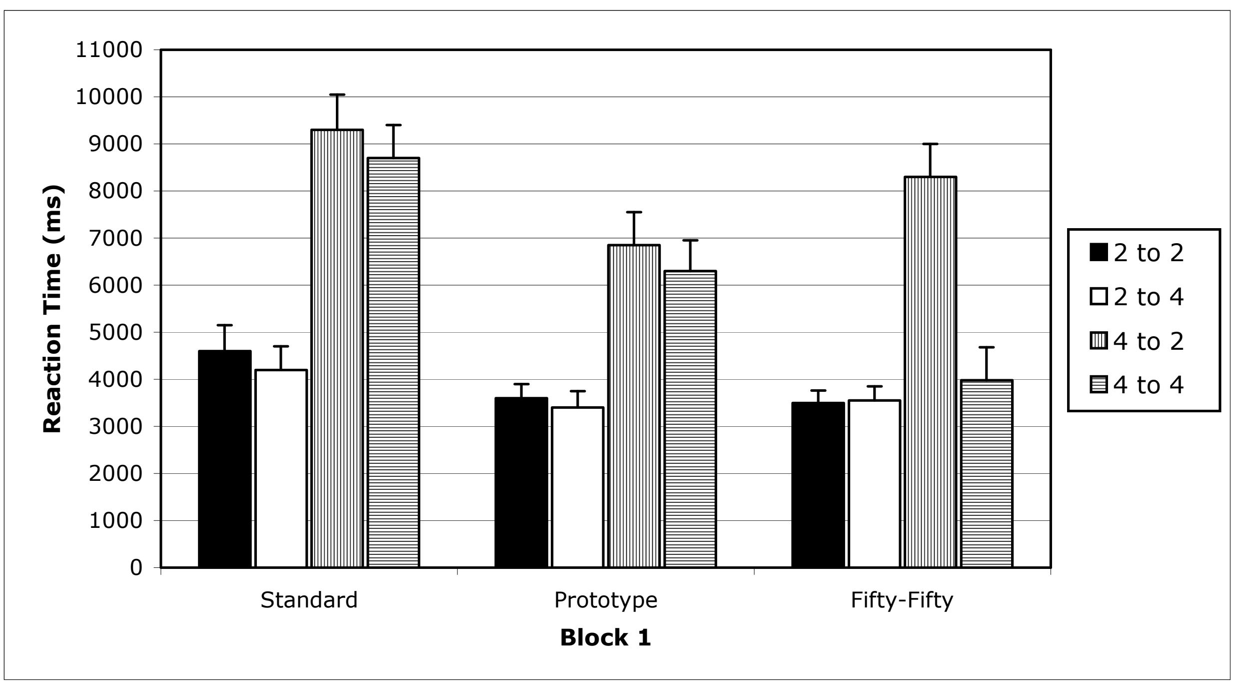
Item learning for the meaningful group in Experiment 5

Figure 23



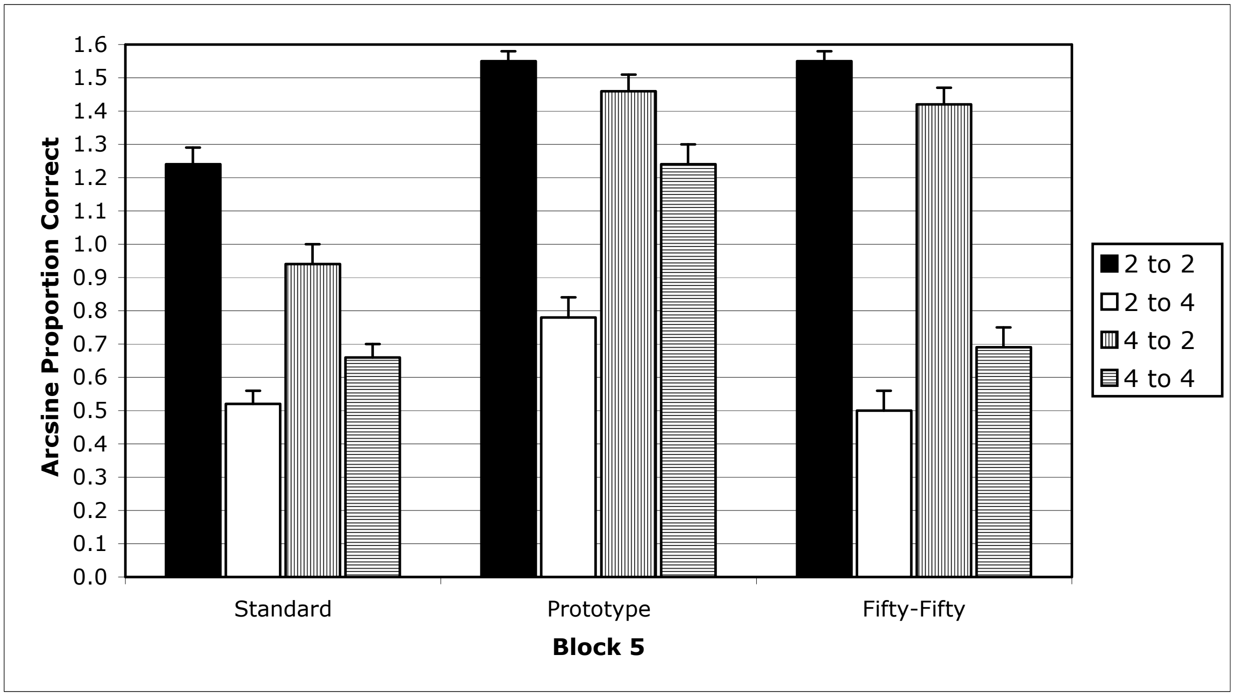
Item response times for the artificial group in Experiment 5

Figure 24



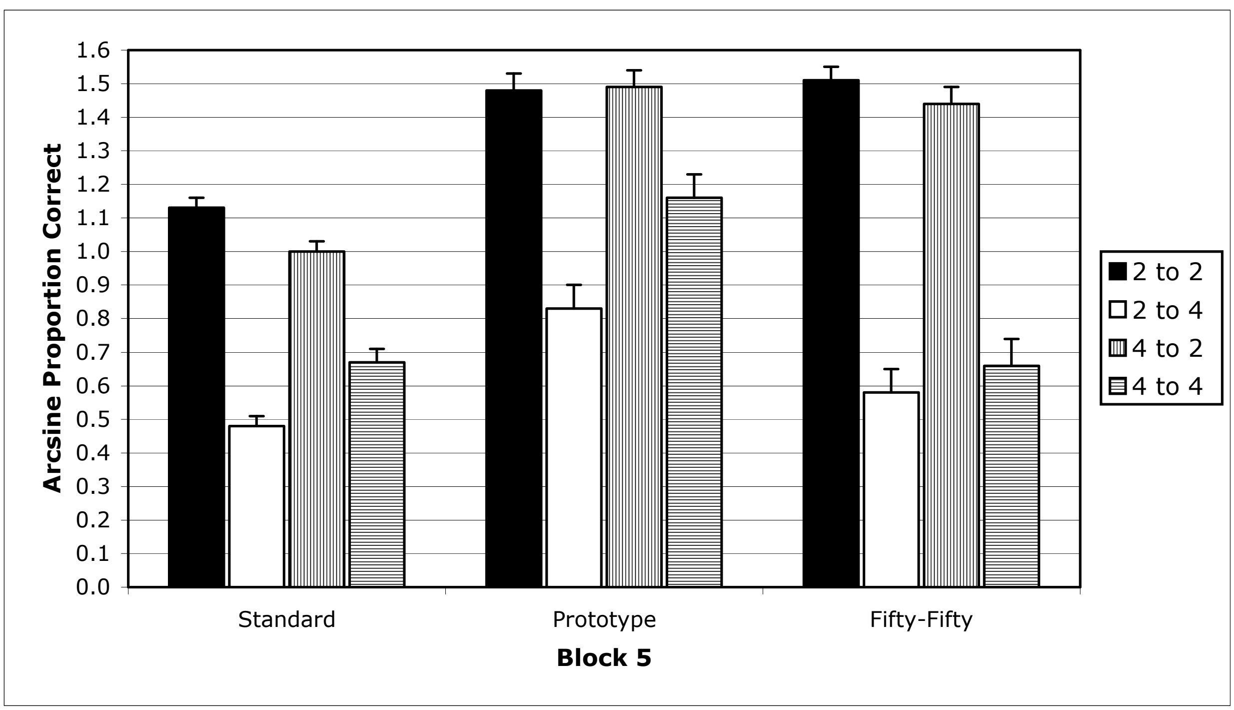
Item response times for the meaningful group in Experiment 5

Figure 25



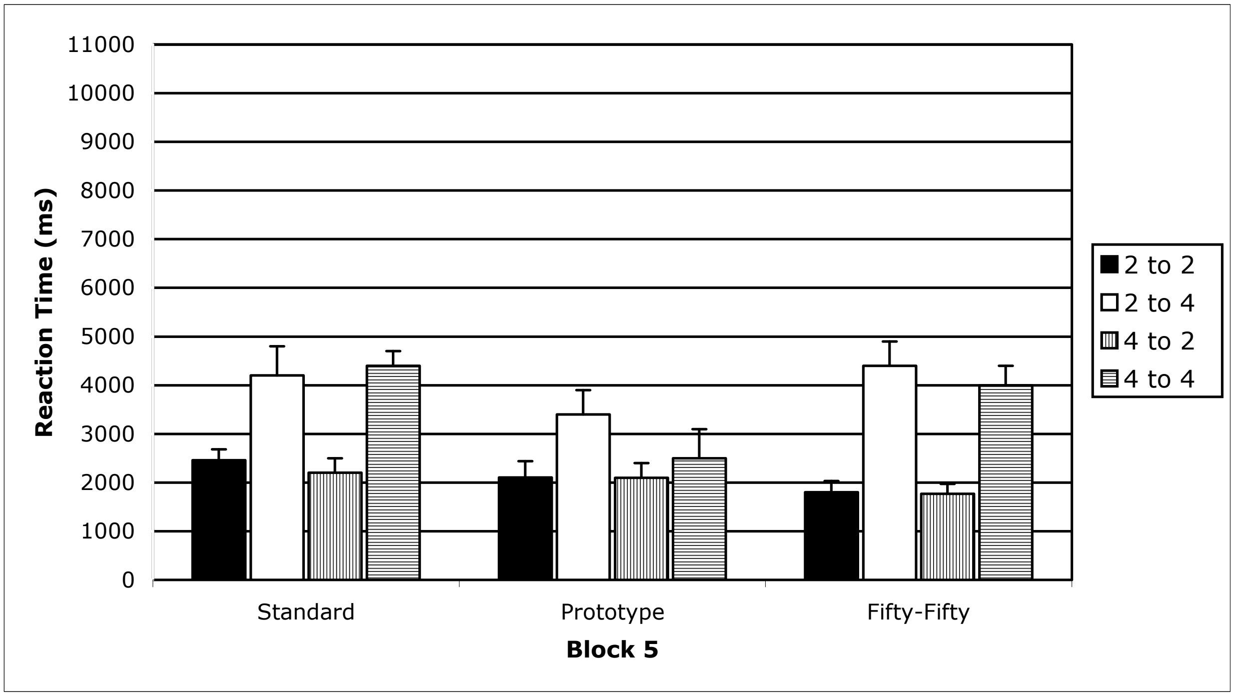
Item learning for the artificial group in Experiment 5

Figure 26



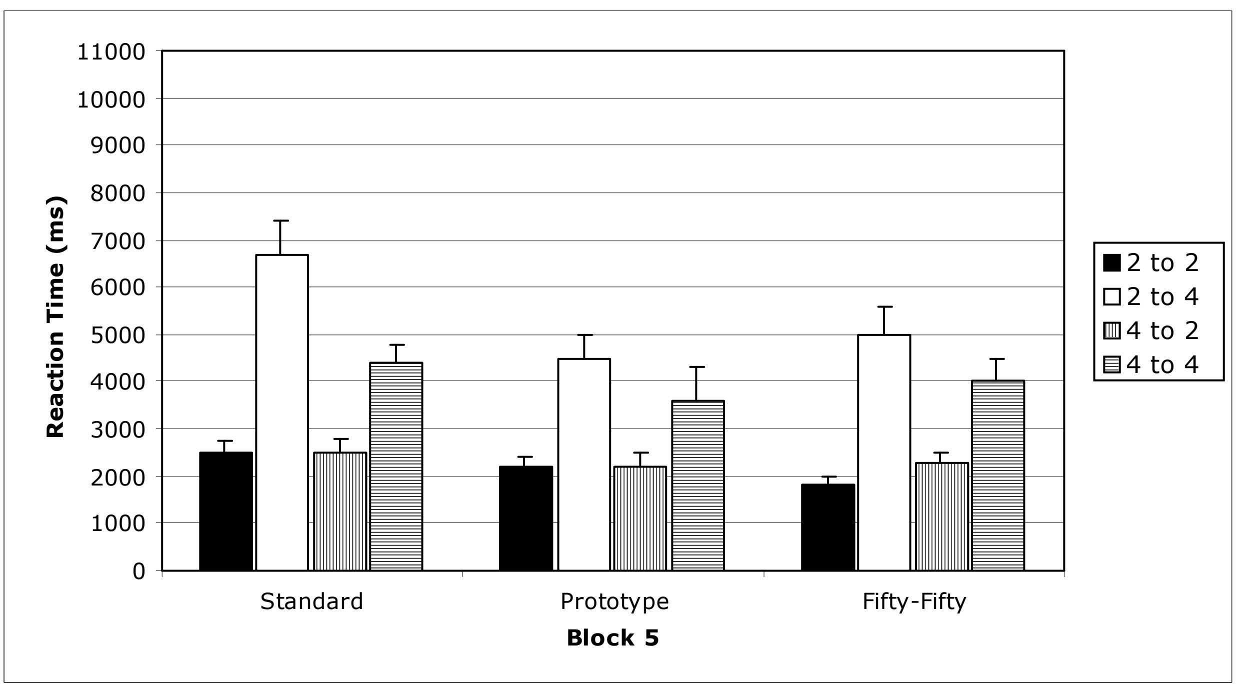
Item learning for the meaningful group in Experiment 5

Figure 27



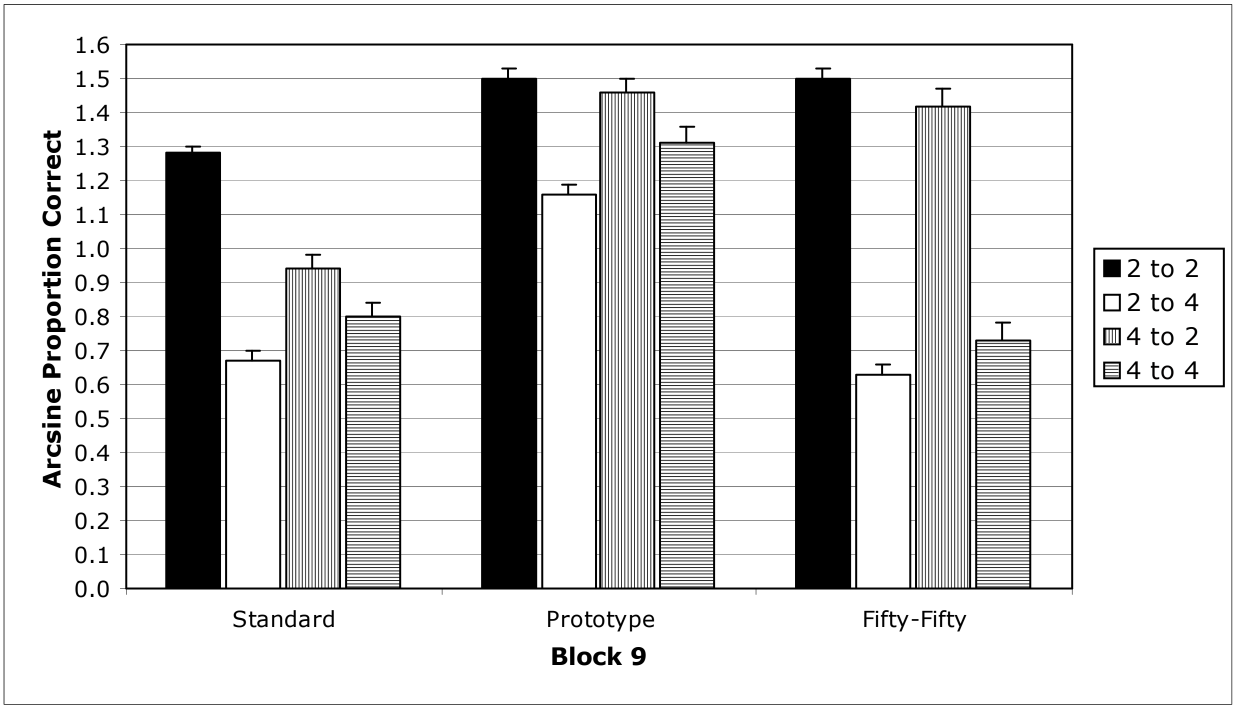
Item response times for the artificial group in Experiment 5

Figure 28



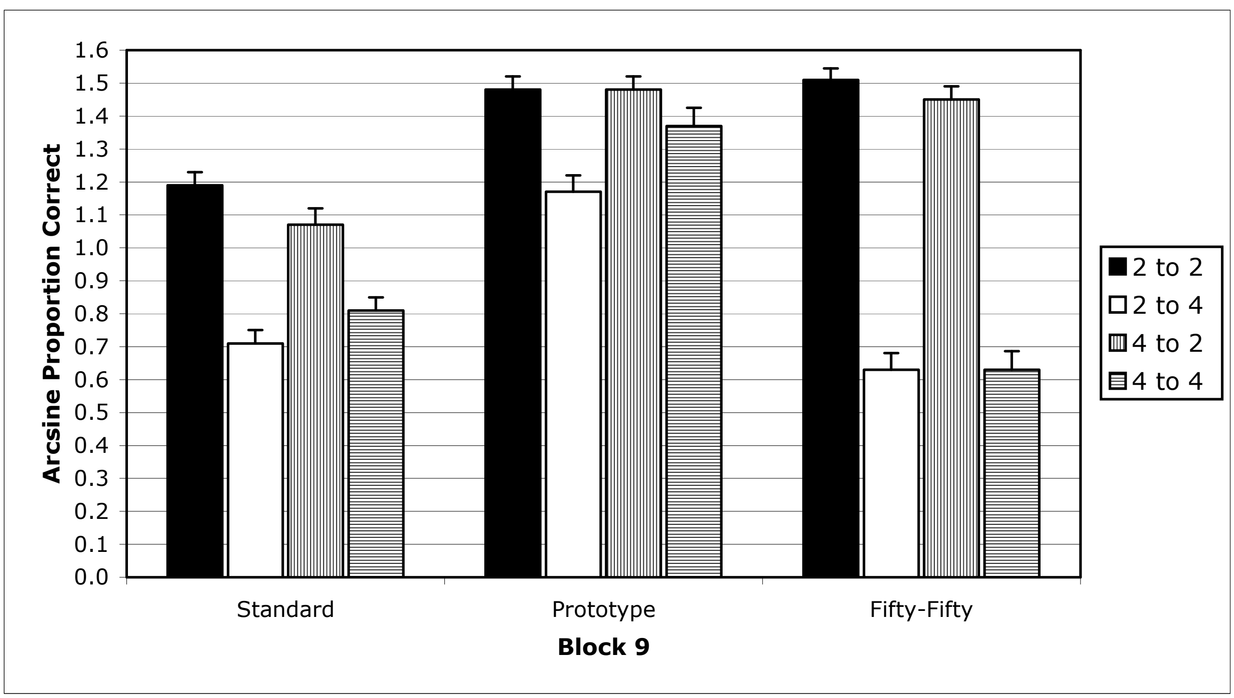
Item response times for the meaningful group in Experiment 5

Figure 29.



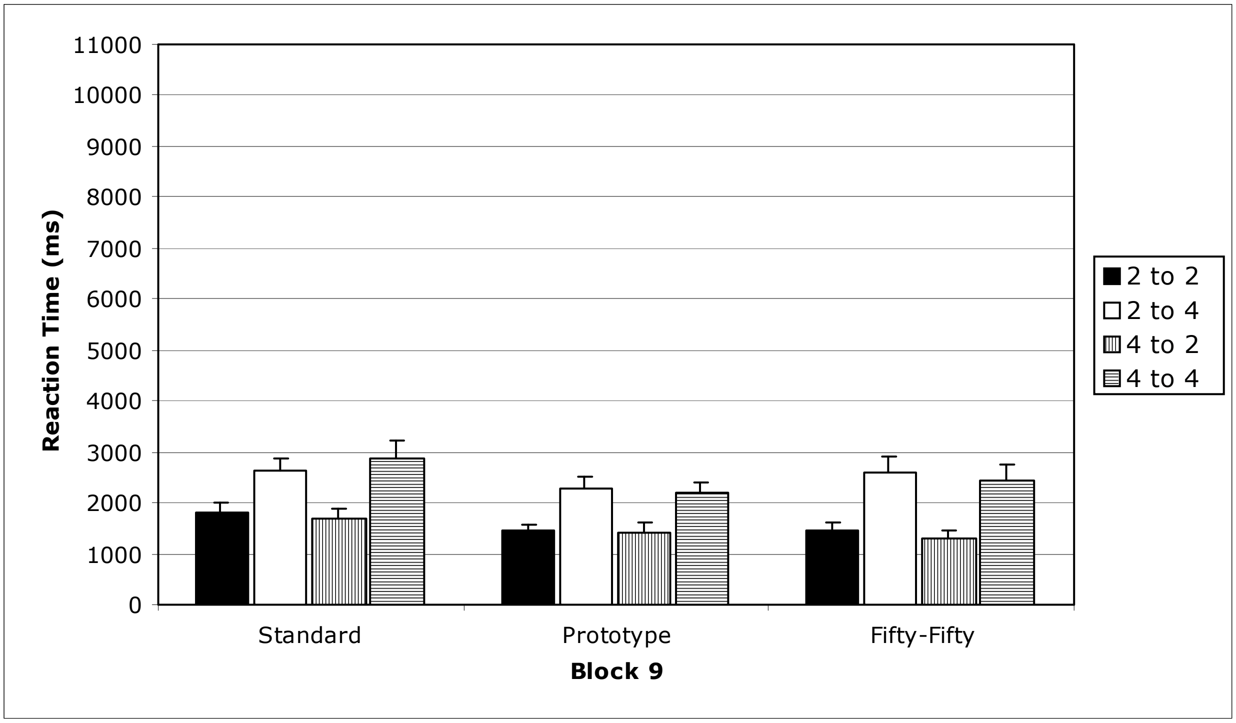
Item learning for the artificial group in Experiment 5

Figure 30.



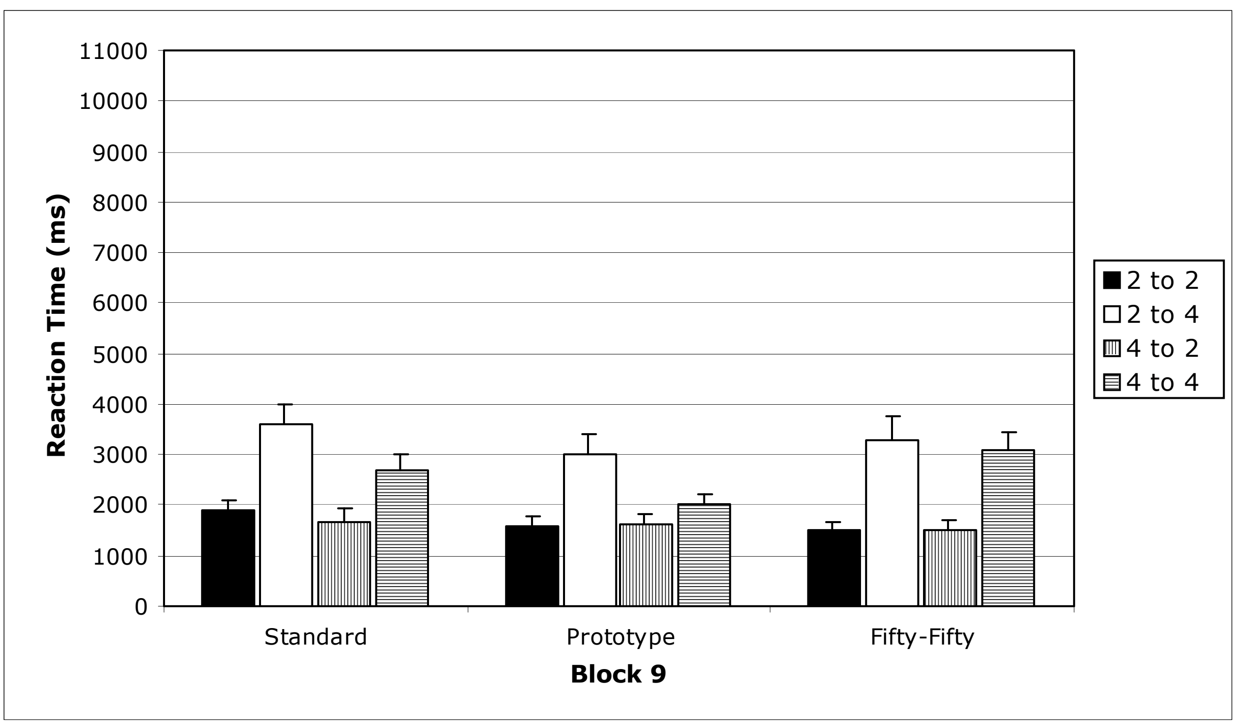
Item learning for the meaningful group in Experiment 5

Figure 31.



Item response times for the artificial group in Experiment 5

Figure 32.



Item response times for the meaningful group in Experiment 5

6.2.1 Item comparisons for artificial and meaningful groups

The following analyses compare group differences when learning standard, prototype, and fifty-fifty items at blocks 1, 5, and 9. As prototype items are congruent with participant's prior expectations one would expect enhanced performance on part of the meaningful group for these items. However, if abstract dimensions interact adversely with prior experiences, weaker performance (or equal performance) is expected on part of the meaningful group for these items.

6.2.1.1 Accuracy

The first sets of analyses compare artificial and meaningful label groups for prototype, fifty-fifty, and standard items, on blocks 1, 5, and 9 for superordinate level categories.

First, when comparing the artificial superordinate-superordinate block 1 condition with the meaningful superordinate-superordinate block 1 condition, results showed a main effect of item type (F (2, 108) = 82.77, $p < .001$). However, the interaction between item type and group was statistically non-significant (F (2, 108) = .26, $p > .77$), as was the main effect of group (F (1, 54) = .66, $p > .42$).

Second, when comparing the artificial superordinate-superordinate block 5 condition with the meaningful superordinate-superordinate block 5 condition, results showed a main effect of item type (F (2, 108) = 65.32, $p < .001$). The interaction between item type and group was statistically non-significant (F (2, 108) = .57, $p > .57$) as was the main effect of group (F (1, 54) = 2.34, $p > .13$).

Finally, when comparing the artificial superordinate-superordinate block 9 condition with the meaningful superordinate-superordinate block 9 condition, results showed a main effect of item type (F (2, 108) = 65.32, $p < .001$). The interaction between item type and group was statistically non-significant (F (2, 108) = .57, $p > .57$), as was the main effect of group (F (1, 54) = 2.35, $p > .13$).

In sum, the artificial and meaningful labels did not differ in performance on prototype, standard, and fifty-fifty items at block 1, 5, and 9, for the superordinate-superordinate phase 1 and 2 conditions.

The next sets of analyses compare artificial and meaningful label groups for prototype, fifty-fifty, and standard items, on blocks 1, 5, and 9 for basic level categories.

First, when comparing the artificial basic-basic group with the meaningful basic-basic group for block 1, results showed a main effect of item type (F (2, 108) = 25.78, $p < .001$). The interaction between item type and group was statistically non-significant (F (2, 108) = 1.59, $p > .21$) as was the main effect of group (F (1, 54) = .99, $p > .33$).

Second, when comparing the artificial basic-basic group with the meaningful basic-basic group for block 5, results showed a main effect of item type (F (2, 108) = 32.20, $p < .001$). The interaction between item type and group was statistically non-significant (F (2, 108) = .15, $p > .86$), as was the main effect of group (F (1, 54) = .17, $p > .68$).

Finally, when comparing the artificial basic-basic group with the meaningful basic-basic group for block 9, results showed a main effect of item type (F (2, 108) = 75.34, $p < .001$). The interaction between item type and group was statistically non-significant (F (2, 108) = .96, $p > .36$) as was the main effect of group (F (1, 54) = .02, $p > .90$).

In sum, the artificial and meaningful labels did not differ in performance on prototype, standard, and fifty-fifty items at block 1, 5, and 9, for the basic-basic phase 1 and 2 conditions

The next sets of comparisons explore differences between groups for taxonomic transfer phase 2 categories.

First, focusing on basic-superordinate phase 2 condition for block 5 results showed a main effect of item type (F (2, 108) = 104.83, $p < .001$. The interaction between item type and group was statistically non-significant, F (2, 108) = .19, $p > .83$, as was the main effect of group (F (1, 54) = .47, $p = .49$). Second, focusing on block 9, results showed a main effect of item type (F (2, 108) = 74.09, $p < .001$). The interaction between item type and group was statistically non-significant (F (2, 108) = .77, $p > .47$) as was the main effect of group (F (1, 54) = 1.05, $p > .31$).

Next, focusing on superordinate-basic phase 2 condition for block 5, results showed a main effect of item type (F (2, 108) = 13.60, $p < .001$). The interaction between item type and group was statistically non-significant (F (2, 108) = .45, $p > .64$) as was the main effect of group (F (1, 54) = .34, $p > .56$). Second, focusing on block 9, results showed a main effect of item type (F (2, 108) = 61.80, $p < .001$). The interaction between item type and group was statistically non-significant (F (2, 108) = .11, $p > .90$) as was the main effect of group (F (1, 54) = .08, $p > .77$).

In sum, results failed to reveal reliable differences between groups when categorizing items at either superordinate or basic levels.

6.2.1.2 Reaction Times

The first sets of analyses compare artificial and meaningful label groups for prototype, fifty-fifty, and standard items, on blocks 1, 5, and 9, for superordinate level categories.

First, when comparing the artificial superordinate-superordinate group with the meaningful superordinate-superordinate group at block 1, results showed a main effect of item type (F (2, 108) = 6.43, $p < .01$). The interaction between item type and group was statistically non-significant (F (2, 108) = .83, $p > .44$) as was the main effect of group (F (1, 54) = .45, $p > .51$).

Second, when comparing the artificial superordinate-superordinate group with the meaningful superordinate-superordinate group at block 5, results showed a main effect of item type (F (2, 108) = 7.91, $p < .001$). The interaction between item type and group was statistically non-significant (F (2, 108) = .02, $p > .98$) as was the main effect of group (F (1, 54) = .04, $p = .84$).

Finally when comparing the artificial superordinate-superordinate group with the meaningful superordinate-superordinate group at block 9, results showed a main effect of item type F (2, 108) = 8.23, $p < .001$. The interaction between group and item type was statistically non-significant (F (2, 108) = .05, $p > .95$) as was the main effect of group (F (1, 54) = .41, $p > .53$).

The first sets of analyses compare artificial and meaningful label groups for prototype, fifty-fifty, and standard items, on blocks 1, 5, and 9, for basic level categories.

First, when comparing the artificial basic-basic phase 1 condition with the meaningful basic-basic phase 1 conditions, results showed a main effect of item type (F

(2, 108) = 5.27, $p < .01$) and a significant main effect of group (F (1, 54) = 11.21, $p < .001$). The interaction between item type and group was statistically non-significant (F (2, 108) = .78, $p > .45$). The *artificial group* was faster than meaningful group when prototype (t (54) = 10.73, $p < .002$), and fifty-fifty, items (t (54) = 5.32, $p < .03$) were classified at block 1, and when standard items were classified at block 1 (t (54) = 10.14, $p < .002$), 2 (t (54) = 5.56, $p < .01$) and 4 (t (54) = 3.95, $p < .02$).

Second, when comparing the artificial basic-basic phase 2 condition with the meaningful basic-basic phase 2 condition at block 5 results showed a significant main effect of item type (F (2, 108) = 5.28, $p < .03$). The interaction between item type and group was statistically non-significant (F (2, 108) = 1.88, $p > .16$) as was the main effect of group (F (1, 54) = .05, $p > .83$).

Finally, when comparing the artificial basic-basic phase 2 condition with the meaningful basic-basic phase 2 condition at block 9, results showed a significant main effect of item type (F (2, 108) = 5.59, $p < .005$). The interaction between item type and group was statistically non-significant (F (2, 108) = 1.24, $p > .29$) as was the main effect of group (F (1, 54) = .01, $p > .98$).

In sum, no statistical differences were found between groups for items when categorizing at superordinate levels. However, faster performance was found for the artificial label group when processing standard, prototype and fifty-fifty items for basic level categories. These findings would suggest that mapping between item information and the category is poorer for the meaningful group during initial stages of learning.

6.2.2 Item comparisons for individual groups

Analyses for Experiment 4 revealed accuracy differences between the three items, but minimal differences in reaction times. This failure to find reaction time differences may have resulted from performance asymptote by block 9. Introducing prototype and fifty-fifty items into block 1 and 5 has the advantage of examining this possibility. An additional advantage follows from closer examination of transfer patterns for prototype and fifty-fifty items.

6.2.2.1 Artificial label group

Means and mean square errors for accuracy are depicted in figures 21, 25, and 29 and for response times in figures 23, 27, and 31. Because there were no logical differences between superordinate-superordinate and superordinate-basic groups or between the basic-superordinate and the basic-basic groups at block 1, these groups were combined for present analyses. Accuracy data are explored first, followed by response times.

6.2.2.2 Accuracy for block 1

As can been seen in figure 21, prototype items were classified better than standard items in the superordinate condition ($\underline{t}$ (27) = 5.32, $\underline{p}$ < .001) and the basic condition ($\underline{t}$ (27) = 3.81, $\underline{p}$ < .001). However, no statistical differences were found between prototype and fifty-fifty item in the superordinate condition ($\underline{t}$ (27) = .43, $\underline{p}$ > .67). Finally, fifty-fifty items were classified better than standard items in the superordinate condition ($\underline{t}$ (27) = 8.68, $\underline{p}$ < .001).

These findings are consistent with the previous experiment and suggest that participants categorize based on the number of statistical values that correctly predict the category.

6.2.2.3 Accuracy for block 5

As can be seen in figure 25, prototype items were classified better than standard items in both the superordinate-superordinate phase 2 (t (27) = 5.08, $p < .001$) and the basic-superordinate phase 2 condition (t (27) = 11.42, $p < .001$). Moreover, prototype items were classified better than standard items in the basic-basic phase 2 condition (t (27) = 8.03, $p < .001$) and the superordinate-basic levels phase 2 condition (t (27) = 3.70, $p < .001$). However, no differences were found between prototype and fifty-fifty items in either the superordinate-superordinate phase 2 condition (t (27) = .01, $p > .99$) or the basic-superordinate levels phase 2 condition, (t (27) = .68, $p > .51$).

Finally, fifty-fifty items were classified better than standard items in the superordinate-superordinate phase 2 (t (27) = 5.69, $p < .001$) and the basic-superordinate levels phase 2 conditions (t (27) = 8.78, $p < .001$).

The failure to find differences between prototype and fifty-fifty items immediately following taxonomic transfer from the basic to superordinate level is important. As noted earlier, given an effect of taxonomy, these items would most likely differ at the point of taxonomic transfer. Particularly in the basic-super, because participants have just been taught categories that corresponded to prototype items better than to fifty-fifty items.

6.2.2.4 Accuracy for block 9

As can be seen in figure 29, prototype items were classified better than standard items in both the superordinate-superordinate phase 2 (t (27) = 4.23, $p < .001$) and basic-superordinate levels phase 2 conditions (t (27) = 8.11, $p < .001$). Prototype items were also classified better than standard items at both the basic-basic phase 2 (t (27) = 9.68, p

< .001) and the superordinate-basic phase 2 conditions, (t (27) = 9.71, $p < .001$). However, no differences were found between prototype and fifty-fifty items in either the superordinate-superordinate phase 2 (t (27) = .05, $p > .96$) or the basic-superordinate phase 2 conditions, t (27) = 1.45, $p > .16$.

Finally, fifty-fifty items were classified better than standard items in both the superordinate-superordinate (t (27) = 4.20, $p < .003$) phase 2 conditon, and the basic-

6.2.2.5 Reaction times for block 1

As can be seen in figure 23, Prototypes were classified faster than standard items in the basic condition (t (27) = 4.02, $p < .001$). However no differences were found between these items in the superordinate condition, (t (27) = 1.48, $p > .15$). Moreover, no differences were found between prototype and fifty-fifty items in the superordinate condition (t (27) = .32, $p > .75$). Finally, no statistical differences were found between fifty-fifty items and standard items in the superordinate condition (t (27) = 1.75, $p > .09$).

In sum, the only reaction time differences found occurred at the basic level with prototype items being classified faster than standard items.

6.2.2.6 Reaction Times for block 5

As can be seen in figure 27, no differences were found between prototype and standard items in either the superordinate-superordinate phase 2 condition (t (27) = 1.08, $p > .29$) or the basic superordinate levels phase 2 condition (t (27) = .18, $p > .86$). When considering basic level comparisons for prototype and standard items, prototype items were classified faster than standard items in the basic-basic phase 2 condition (t (27) = 3.21, $p < .003$) but not in the superordinate-basic phase 2 condition (t (27) = 1.38, $p > .18$). Furthermore, no significant differences were found between prototype and fifty-fifty items in the superordinate-superordinate phase 2 (t (27) = .83, p = .42) and the basic-superordinate phase 2 condition (t (27) = 1.97, p > .06).

Finally, fifty-fifty items were classified faster than standard items in the superordinate-superordinate phase 2 (t (27) = 3.81, p < .001) and the basic-superordinate phase 2 conditions (t (27) = 2.38, p < .03).

In sum, prototype items were classified faster than standard items at the basic-basic phase 2 condition. Moreover, fifty-fifty items were classified faster than standard items in the superordinate-superordinate phase 2, and the basic-superordinate phase 2 conditions.

6.2.2.7 Reaction times for block 9

As can be seen in figure 31, prototype items were classified faster than standard items in the both the superordinate-superordinate phase 2 (t (27) = 2.22, p < .04) condition and the basic-superordinate phase 2 condition (t (27) = 2.15, p < .04). However, no significant differences were found between prototype and standard items in both the basic-basic phase 2 (t (27) = 1.69, p > .10) and the superordinate-basic phase 2 conditions (t (27) = 1.48, p > .15). Furthermore, no differences were found between prototype and fifty-fifty items in either the superordinate-superordinate phase 2 (t (27) = .03, p > .98) or basic-superordinate phase 2 conditions (t (27) = .66, p > .52).

Finally, fifty-fifty items were classified faster than standard items in both the superordinate-superordinate condition (t (27) = 2.50, p < .03) and the basic-superordinate condition, t (27) = 3.66, p < .001.

In sum, no response time differences were found between prototype and fifty-fifty items, however response times for these items were faster than for standard items.

6.2.3 Meaningful label group

Means and mean square errors for accuracy are depicted in figures 22, 26, and 30, and for response time in figures 24, 28, and 32. Because there were no logical differences between superordinate-superordinate and superordinate-basic groups or between the basic-superordinate and the basic-basic groups these groups in block 1, these groups were combined for analysis. Accuracy data are explored first, followed by response times.

6.2.3.1 Accuracy for Block 1

As can be seen in figure 22, prototype items were classified better than standard items in both the superordinate ($\underline{t}$ (27) = 10.97, $\underline{p} < .001$) and basic conditions ($\underline{t}$ (27) = 5.62, $\underline{p} < .001$). However, no statistical differences were found between prototype and fifty-fifty items in the superordinate condition ($\underline{t}$ (27) = .85, $\underline{p} > .40$). Finally, fifty-fifty items were classified better than standard items in the superordinate condition ($\underline{t}$ (27) = 7.34, $\underline{p} < .001$).

In sum, prototype and fifty-fifty items were classified better than standard items, however no differences were found between fifty-fifty items and prototype items.

Thus, findings do not support the idea that participants viewed fifty-fifty items as odd otherwise performance for these items would have been poorer than for other items. Instead, findings suggest that participants classified based on the number of dimensional values that correctly predicted the category. Other explanations are explored in the discussion for this experiment.

6.2.3.2 Accuracy for Block 5

As can be seen in figure 26, prototype items were classified better than standard items in the superordinate-superordinate phase 2 condition (t (27) = 6.91, p < .001) and the basic-superordinate phase 2 condition (t (27) = 10.37, p < .001). Furthermore, prototypes were also classified better than standard items in both the basic-basic phase 2 condition (t (27) = 6.39, p < .001) and the superordinate-basic phase 2 condition (t (27) = 4.14, p < .001). However, no differences were found between prototype and fifty-fifty items in either the superordinate-superordinate phase 2 condition (t (27) = 1.00, p > .33) or the basic-superordinate phase 2 condition (t (27) = .92, p > .37).

Finally, fifty-fifty items were classified better than standard items in both the superordinate-superordinate phase 2 condition (t (27) = 4.44, p < .001) and the basic-superordinate phase 2 condition (t (27) = 8.66, p < .001).

In sum, findings for block 5 replicate those of block 1, prototype and fifty-fifty items were classified better than standard items, however no differences were found between fifty-fifty items and prototype items. An important finding here is the failure to find differences between prototype and fifty-fifty items immediately following taxonomic transfer from basic to superordinate level. As noted earlier, taxonomic transfer is the point at which differences between these items was most likely to occur.

6.2.3.3 Accuracy for block 9

As can be seen in figure 30, prototype items were classified better than standard items in both the superordinate-superordinate phase 2 condition (t (27) = 10.97, p < .001) and the basic-superordinate phase 2 condition (t (27) = 7.63, p < .001). Moreover, prototypes were classified better than standard items (t (27) = 9.13, p < .001) in both the basic-basic phase 2 condition (t (27) = 6.80, p < .001) and the superordinate-basic phase 2 condition. However, no differences were found between prototype and fifty-fifty items when classified in either the superordinate-superordinate phase 2 condition (t (27) = .85, p > .40) or the basic-superordinate phase 2 condition (t (27) = .59, p > .56).

Finally, fifty-fifty items were classified better than standard items in both the superordinate-superordinate phase 2 condition (t (27) = 7.51, p < .001) and the basic-superordinate phase 2 condition (t (27) = 6.89, p < .001).

6.2.3.4 Reaction times for block 1

As can be seen in figure 24, prototypes items were classified faster than standard items in both the superordinate condition (t (27) = 3.52, p < .002) and the basic condition (t (27) = 3.05, p < .02). However, no differences were found between prototype and fifty-fifty items in the superordinate condition (t (27) = .05, p > .96). Finally, fifty-fifty items were classified faster than standard items in the superordinate condition (t (27) = 3.99, p < .001).

In sum, the finding of faster reaction time for prototype and fifty-fifty items over standard items coupled with the failure to find differences between prototype and fifty-fifty items suggests that participants classified based on the number of dimensional values belonging to categories.

6.2.3.5 Reaction times for block 5

As can be seen in figure 28, no reaction time differences were found between prototype and standard items in either the superordinate-superordinate phase 2 (t (27) = 1.95, p > .06) and basic-superordinate phase 2 conditions (t (27) = 1.29, p > .21). When comparing differences between prototype and standard items at the basic levels, prototypes items were classified faster than standard items in both the superordinate-basic phase 2 condition (t (27) = 2.35, $p < .05$) and the basic-basic phase 2 condition (t (27) = 4.21, $p < .001$). No differences were found between prototype and fifty-fifty items when classified in either the superordinate-superordinate phase 2 condition (t (27) = 1.20, $p >$ .06) or the basic-superordinate phase 2 condition (t (27) = .35, $p > .73$).

Finally, fifty-fifty items were classified faster than standard items (t (27) = 5.29, p < .001) in the superordinate-superordinate phase 2 condition, but not in the basic-superordinate phase 2 condition (t (27) = .50, $p > .62$).

6.2.3.6 Reaction times for block 9

As can be seen in the figure 32, prototype items were classified faster than standard items in the superordinate-superordinate phase 2 condition (t (27) = 3.18, $p <$ .004) but not in the basic-superordinate phase 2 condition (t (27) = .38, $p > .73$). Prototype items were also classified faster than standard items in both the basic-basic phase 2 condition (t (27) = 2.23, $p < .03$) and the superordinate-basic phase 2 condition (t (27) = 2.76, $p < .01$). However, no differences were found between prototype and fifty-fifty items when classified in either the superordinate-superordinate phase 2 condition (t (27) = .49, $p > .63$) or the basic-superordinate phase 2 condition (t (27) = .77, $p > .45$).

Finally, fifty-fifty items were classified faster than standard items in the superordinate-superordinate phase 2 condition (t (27) = 4.72, $p < .001$) but not in the basic-superordinate phase 2 condition (t (27) = 1.35, $p > .19$).

In comparisons to standard items, participants were generally faster processing prototype and fifty-fifty items. The one exception was the failure to find a difference between these items for the basic-superordinate phase 2 condition. It would seem that standard items benefited more from basic level priming than other items. Indeed, reaction times for these items were faster when transferring from basic to superordinate levels than when transferring from superordinate to superordinate levels. This finding may be limited to data set, as similar findings were not found in Experiment 4.

6.2.4 Differences between taxonomic levels

The next sets of analyses examine mean response time differences between superordinate and basic level categories for artificial and meaningful groups. These analyses are important for differentiating the influence of prior expectations, dimensions, and item structure, on categorization. For example, slower responses on part of the meaningful group for prototype items would suggest that abstract dimensions negatively impacted performance. This is because the primary factor affecting performance for prototype items is abstract dimensions (structure for prototype items was held constant between taxonomic levels). However, slower responses to standard items on part of the meaningful group would suggest that both abstract dimensions and incongruent dimensional values negatively affected meaningful group performance (these items have both abstract dimensions and incongruent dimensional-values). Finally, slower between taxonomic level responses on part of the meaningful group when comparing standard and prototype items on would suggest that incongruent dimensional values are the primary factor affecting performance. The first set of analyses compares mean differences between basic-basic and superordinate-superordinate groups. The second set of analyses examines group mean differences between basic-superordinate and superordinate-basic phase 2 conditions.

Results were inconclusive when comparing differences between taxonomic levels for meaningful and artificial groups in Experiment 4. Findings showed that the artificial group was faster processing standard items, however no differences were found between groups for prototype items. Thus, it was difficult to determine with any certainty to what extent abstract or incongruent dimensional values contributed to findings. One reason for failing to find an effect of abstract dimensions for prototype items may follow from items' characteristics being less surprising to the meaningful group by block 9. Given this possibility one would expect greater differences between groups when processing prototype items at block 1. For similar reasons one might also expect that given an effect of incongruent dimensional values, differences between taxonomic levels when comparing prototype and standard items would also be greater for the meaningful label at block 1.

First, when comparing differences between basic-basic and superordinate-superordinate groups at block 1, mean differences were smaller for the artificial label when categorizing standard (t (54) = 2.56, $p < .01$) and prototype items (t (54) = 3.05, $p < .01$). No statistical differences were found between groups when categorizing items on blocks 5 ($p > .43$) and 9 ($p > .74$). These findings support the idea that abstract dimensions interfere with the meaningful groups mapping of item information. Second, comparing differences between basic-superordinate and superordinate-basic phase 2 condition, differences were smaller for the artificial group when categorizing standard items at block 5 (t (54) = 2.98, $p < .03$). All other comparisons were statistically non-significant ($p > .82$).

Finally, the next comparison explores the idea that incongruent dimensional values interfere with mapping of item information. First, focusing on block 1, although observable mean differences between standard and prototype items were greater for the meaningful group (M = 1579) than the artificial group (M = 663) findings were statistically non-significant (t (54) = 1.02, p > .74). Comparisons were also statistically non-significant, at blocks 5 (t (54) = .98, p > .92) and 9 (t (54) = .88, p > .94). Thus, at least statistically an effect of dimensional values goes unsupported.

An important question asked in Experiment 4 centered on the extent to which abstract dimensions and incongruent dimensional values affected the performance of the meaningful group. In that experiment the findings on this question were inconclusive. However, this experiment suggests that abstract dimensions negatively affected response times of meaningful group participants. Between taxonomic level comparisons showed that performance for the meaningful group was poorer on prototype items (than the artificial group). This finding was found only during initial training, suggesting the meaningful participants found this information most problematic when first encountered and at the point when they found them most surprising. Indeed, as training progressed the effect of the label appeared to diminish. Although meaningful participants were faster responding to items as training progressed, their performance still did not exceed artificial participants in terms of either response times or accuracy. This would suggest that the meaningful participants initially tried to map abstract item information to the known category, but failing to do so opted for an alternative strategy. As the artificial and meaningful conditions were quite similar by the end of training, rote learning may have been one strategy opted for by the meaningful group.

Other research demonstrating similar findings used information outside the domain of the category (e.g., Heit & Bott 2000; Heit et al. 2004). However, the primary factor driving present findings was abstract dimensions that were unfamiliar within their domain. Overall, the dimensions in the present experiment map cleanly to the category of musical instruments. However, because participants had only partial knowledge of these dimensions they found it difficult to make clear connections between the dimensions and the known category. It is also important to consider an influence of incongruent dimensional values (i.e., note that values are incongruent in that participants may find that 50 kg. *flutes* contradict prior experiences with *flutes*). In comparison to prototype items, participants clearly found standard items more difficult to categorize. This may follow from standard items having dimensional values that at times contradict participants' prior experiences (e.g., 50 kg. *flutes*), or because they included features predicting a different category. Even though these features were not encountered for every item, even a few encounters may have caused participants to reconsider their categorizations.

Another finding was consistent with that of the previous experiment. No differences were found between prototype and fifty-fifty items. This would suggest that participants were not making classification judgments based on prior experience. Otherwise, participants would have seen fifty-fifty items as either half-flute/half saxophone or half drum/half bell and poorer performance would have been the outcome. Given that all dimensional values for both prototype and fifty-fifty items belonged to their superordinate category, participants may have made their classification decisions based on the number of dimensional-values correctly predicting the category. Another contributing factor to participants' inability to distinguish these items may follow from the relationship between dimensional values and categories. Values for prototype and fifty-fifty items are similar, thus unless participants critically evaluated the differences between items they may have missed subtle distinctions (however, in the basic-superordinate condition participants spent four blocks differentiating the dimensional values and there was still no affect suggesting that even though they knew the values were different they generalized anyway, see below). Moreover, dimensional values for prototype and fifty-fifty items are highly predictive of similar categories (e.g., dimensional values for prototype and fifty-fifty items are predictive of *flute* and *saxophone*). This would suggest that increasing the space between dimensional values and items would result in greater ability to distinguish prototype and fifty-fifty items.

The observation that participants made their classification judgments based on the similarity and distinction between dimensional values and items may also be suggestive of a basic classification principle. When category information is unclear people may resort to making classification decisions based on how attributes of items are similar and distinctive. Certainly, it seems the meaningful group attempted to make initial categorizations based on prior knowledge. However, when it became apparent that relying on prior knowledge alone was inadequate they relied on similarity judgments. Evidence not presented in the body of this paper but available upon request further supports this observation. Most mistakes made for prototype and fifty-fifty items were made within the superordinate categories. That is, if an item was a flute or saxophone participants were most likely to miss-categorize the item as either saxophone or flute than either *drum* or *bell* (and vice versa). However, mistakes for standard items were as likely to occur outside the superordinate as within. That is, in comparison to prototype and fifty-fifty items if the item was a *flute*, participants were more likely to classify the item as *drum* or a *saxophone*.

Finally, generalization also appears to be a factor in learning taxonomies. The basic taxonomic affect found in Experiments 1-4 was also found in this experiment, and is consistent with participants noticing the superordinate category while learning the basic level categories (i.e., they seem to be able to generalize from the basic level to the superordinate, but seem not to analyze the superordinate into its basic-level constituents). Related evidence occurs in the current experiment in that no differences were found between fifty-fifty items and prototype items following transfer from basic to superordinate levels. That is, if participants have learned that the flute category and the saxophone category are very similar, and indeed are both part of a higher level category, then the fact that the fifty-fifty item is a blend of the two lower level categories may not adversely affect its membership in the higher level category. Thus, generalization may be an important factor influencing classification judgments.

In sum, this experiment produced several interesting findings. First, it was found that one factor affecting meaningful group performance was the abstract quality of our dimensions. Furthermore, meaningful participants found abstract dimensions more surprising during initial stages of learning. Second, when item information is abstract and prior knowledge is uncertain or absent, similarity and generalization may be default strategies used for classification.

Chapter VII

General Discussion

The impetus for this study was to examine a taxonomic learning paradigm which to date had largely been ignored in the category literature. When people learn objects they often do so in the context of complex learning environments. Object attributes must be extracted, compared and learned in the presence of multiple domains and categories. Thus far, most category learning paradigms have only included two category structures. However, two--category learning fails to capture complex learning environments, particularly with respect to taxonomic structures. To this end, this study is one of the few studies to focus on learning of four--category hierarchical structures. The first result of interest was found in Experiment 1 and supported previous research (e.g., see Murphy, 2002) by demonstrating that participants were indeed able to learn four-- category hierarchical structures. In addition there were several interesting findings of not found in previous research (e.g., see Murphy & Smith, 1982) focusing on learning within taxonomies. For one, transferring from four categories in phase 1 to new higher level categories in phase 2 produced performance advantages not evident when transferring from two categories in phase 1 to four categories in phase 2. Participants in the former group generally found learning of materials easier than participants in the latter group. Moreover, taxonomic transfer affects were greater when transferring from four to two categories. This result led to the conclusion that participants transferring from four to two categories found generalization easier than participants in other conditions. That is, participants learning basic level first seemed to noticed the superordinate category while learning the basic level categories and as a result were better able to generalize to the superordinate levels (further discussion of generalization follows below). Generally, however, very little affects of taxonomic transfer were found in Experiment 1. This was evident in that taxonomic transfer effects in the form of accuracy were not found.

The failure to find pervasive taxonomic transfer effects led to the introduction of prior knowledge in the form of category labels in Experiment 2. Prior knowledge in the form of category labels has been shown to guide learning by providing an explanation for the properties and structure of categories (Kaplan & Murphy, 2000). Although introducing meaningful labels was expected to facilitate learning category membership of the items this outcome failed to emerge. Furthermore, the presence of meaningful labels should have made the taxonomic structure obvious. However, there was once again little evidence of taxonomic transfer, though the generalization effects driven by basic level learning found in Experiment 1 were once again confirmed.

In Experiment 3, meaningful and artificial labels groups were compared directly. Furthermore, instructions were introduced that defined how one feature was related to another, and that clarified the taxonomic relations. The expectation was that in comparison to when the taxonomic label was unknown, knowledge of feature relations would boost the manipulation of knowledge associated with the meaningful label. Evidence for this prediction failed to emerge. Indeed, the surprising result was that reaction times were much slower for meaningful participants, particularly when learning basic level categories. This result was explained by suggesting that prior experiences affected learning differently depending on how abstract feature information interacted with the specificity of the taxonomic level. That is, slower reaction times for meaningful participants occurred because some item dimensions were abstract and difficult to attach to prior experiences. Moreover, some dimensional values were inconsistent with their prior experiences of the item (e.g., 55 kg. flute). Meaningful participants therefore needed additional time to consider the relationship between the item and the prior knowledge associated with a taxonomic level. Response times were much slower for meaningful participants when classifying at the basic level because the basic level cues specific experiences for the instrument, whereas the superordinate level cues more general information (see experiment 3 for further explanation).

In Experiment 4, the idea that prior experiences may have negatively affected performance for the meaningful group was further explored by introducing two additional items. The additional items, prototype and fifty-fifty, differed from the standard items used in the previous experiments with respect to the number of features related to the category. All dimensional values for prototype items belonged to their categories. Thus, prototype items were an excellent match at either basic or superordinate levels. Fifty-fifty items were an odd item in that when categorized at the basic level their dimensional values where split equally between the categories. For example, one fifty-fifty item was by its features, half saxophone and half flute. However, when categorized at the superordinate level all dimensional values for fifty-fifty items belong to their category (i.e., although the features indicated half flute and half saxophone, all indicated wind instrument). The inclusion of these items resulted in several predictions. The first prediction suggested that performance for the meaningful group would vary depending on the item classified. Prototypes having all dimensional values belonging to their category and thus consistent with prior expectations should be easiest followed by standard, then fifty-fifty (see introduction Experiment 4 for further explanation). The second prediction focused on why slower response times were found for the meaningful group in Experiment 3. Specifically, to what extent did incongruent dimensional values and abstract dimensions negatively affect response times for the meaningful group? Results for prediction 1 surprisingly revealed no differences between prototype and fifty-fifty items when categorizing at the superordinate level. Outcomes for prediction 2 were somewhat inconclusive in that results only suggested that abstract dimensions negatively affected the meaningful participants' performance.

In Experiment 5, prototype and fifty-fifty items were introduced into blocks 1 and 5. In Experiment 3, greater differences where found between meaningful and artificial groups during the first several blocks of training. With this result in mind it was hypothesized that affects of abstract dimensions and/or incongruent dimensional values would be more apparent during initial training blocks. The results confirmed this suggestion and showed that abstract dimensions resulted in slower response times for the meaningful group. Moreover, the results suggested that similarity of dimensional values and how dimensional values generalize plays an important role in early categorization (see discussion Experiment 5 for further explanation). Finally, the prototype and fifty-fifty items continued to show equivalent performance.

7.1 Taxonomic structures

Although only small positive taxonomic transfer effects were found in this study, other taxonomic effects were large. For one, in comparison to participants exposed to phase 1 conditions, participants exposed to taxonomic transfer conditions were almost always faster processing items. Moreover, meaningful label participants clearly detected differences between taxonomic levels. In comparison to the artificial label participants, meaningful label participants were much slower responding to the basic level than to the superordinate level. Indeed, response time differences between taxonomic levels were very relatively small for the artificial label group but very large for the meaningful label group. This would suggest that meaningful label participants experienced basic and superordinate level differently. Importantly, although effects of taxonomic transfer were found overall they were small, and failed to emerge almost entirely in some conditions (e.g., the superordinate-basic condition). Prior research (Murphy & Smith, 1982; Rosch et al. 1976) may provide some insight into why only small taxonomic transfer effects were found.

Much prior research (see e.g., Murphy & Smith, 1982; Rosch et al. 1976) has strongly supported the basic level as the level of special distinction. In general, people can classify objects at different levels of abstraction. They can categorize an object at the superordinate level (e.g., *wind instrument*), the subordinate level (e.g., *kettle drum*), or an intermediate level (e.g., *flute*). Rosch et al. (1976) established the intermediate basic level as the preferred level for categorization: It is the level at which objects are spontaneously labeled; it is the level for fastest categorization and identification of objects; and it is the place where most feature-based information for categories members are stored. The basic level advantage has since been replicated and extended across domains. For example, in one of the few papers to employ a taxonomic learning paradigm, Murphy and Smith (1982) controlled linguistic factors by using artificial categories and yielded basic level superiority effects similar to those of Rosch et al. (1976). Their findings ruled out the possibility that basic level effects resulted solely from linguistic factors (e.g., saliency, word length, frequency). Most relevant to present results, Murphy and Smith found support for the idea that the basic level category advantage arose as a result of their combination of information and distinctive attributes, however this claim was qualified by the need for attributes to be perceptual in character.

That Murphy and Smith found it necessary for attributes to be perceptual in quality may also explain on some level why meaningful participants found categorization difficult in this study. In this study, abstract item information being largely non-perceptual may have interfered with participants' ability to visualize the category. People have a clear idea of what features are central to the categories of flute and drum. Mention the category *flute*, and many will consider an instrument light in weight, silver in color, high in pitch that is either played in the school orchestra or seen at the local symphony. Because participants had very clear prior ideas and perceptions of the instruments classified in this study they may have found it difficult to translate abstract features into features they knew something about (e.g., translating resonate frequency into pitch). Moreover, individual features rarely define an entire object. Classifying an object as a *bird* when the only knowledge one has of the object is that it has wings would be difficult (planes also have wings). Objects are more like sets of features that form a perceptual whole. Meaningful label participants may have found it very difficult to connect the abstract features used in the present study to one another in a way that resulted in their forming a holistic impression of the instrument. Their inability to correlate features to one another and map these features clearly to the prior expectations (the category label) may have resulted in their having decidedly slower response times (see Experiment 3 and below for further discussion) when classifying at the basic level.

Other evidence points to advantages when learning basic level categories first.

More often than not when participants learned basic level categories first performance advantages in subsequent stages extended further than when they had learned superordinate levels first. Perhaps one reason for this finding is that basic level exposure leads to complex comparisons and deeper examination of the item's structure. Participants exposed to basic level categories are in a position to compare features over a greater number of categories than participants exposed to superordinate categories. These comparisons may lead to a better understanding of high level categories not yet learned and a clearer idea of how items generalize to those categories. Furthermore, because basic level categories are much more similar, comparisons between such categories must be more precise than comparisons between far less similar superordinate categories.

Finally, positive taxonomic transfer effects were by and large found only when transferring from basic to superordinate levels. As noted earlier, this finding is suggestive of generalization. Item structures in this study favor generalization (e.g., the frequent dimensional values are highly similar across the two categories within a superordinate), as do the procedures (e.g., the mapping of categories to response keys). Thus, in some respects it is not surprising that participants mostly preferred a strategy of generalization to critical evaluation. This argument is perhaps stronger for the artificial group (at least initially) as response times for the meaningful group were slower, suggesting at least an attempt at critical evaluation. Nevertheless in the end it would appear that both groups gravitated toward generalization. This tendency occurred in spite of a vast amount of research (see e.g., Murphy, 2000) demonstrating facilitation effects of prior knowledge. The next section examines reasons why prior knowledge effects failed to materialize.

7.2 Prior knowledge

Much prior research has demonstrated that category learning is easier when prior knowledge is consistent with the learned category than when knowledge is incongruent or absent (e.g., Murphy & Allophenna, 1994; Wattenmaker et al. 1996). One common method by which prior knowledge is activated in category learning paradigms is through connection of thematic features that have strong relations to one another and to a category theme. In Murphy and Allopenna's (1994) (see also Ahn, 1991; Murphy & Wisniewski, 1989; Pazzani, 1991; Wattenmaker et al. 1986) study, for instance, participants learned about two contrasting pairs of categories, neutral and integrated. Although neutral and integrated category structures were identical, in that each feature associated with its category occurred in one third of its instances and in none of the instances of the contrasting category, and also from the same domain (e.g., jungle and arctic vehicles) the features comprising the integrated category structures formed a coherent theme. Integrated categories contain features intended to cue participants' prior knowledge. For example, *jungle vehicles* are likely to have wheels and be lightly insulated whereas *arctic vehicles* are more likely to be heavily insulated and drive on glaciers. Participants were never informed of the theme but were expected to identify the theme through learning. In contrast, for the neutral categories, features were thematically unrelated therefore activation of prior knowledge was not expected. For this group, it is improbable that vehicles with a *manual transmission* are more likely than vehicles with an *automatic transmission* to have *radial tires* and *air bags*. The results for this study showed that compared with the participants learning neutral categories, participants learning integrated categories were much faster.

In Murphy and Allopenna's (1994) experiments all of the features were related to the specific category theme. Other research has demonstrated powerful prior knowledge effects even when some, but not all, features were inconsistent with the thematic content of the category. Kaplan and Murphy (2000) introduced category themes similar to that of Murphy and Allopenna. However, instead of prior knowledge relating all relevant features of the category items, every item had only one feature related to a theme (e.g., *jungle* or *arctic vehicle*). Furthermore, these features were spread across items so one item might have drives in jungle, another heavily insulated and so on. Thus, in contrast to previous research, feature connections to the category theme were weak, and participants had to detect the theme by observing how features connected (to that theme) over multiple items. Kaplan and Murphy found that even under these conditions participants learned items much faster when knowledge related features were present than when they where not.

In natural contexts, feature relations can be a powerful means of cueing background knowledge and activating feature correlations. As features do not occur independently of one another, there is a statistical structure in which features co-occur across categories and concepts. There is variation in the degree to which the presence of one feature signals the presence of another. For example, has wings and flies are highly correlated because many types of *birds* that have wings are also likely to fly. In contrast, has whiskers and meows are weakly correlated because things in the world that meow always have whiskers, but there are many types of *animals* that have whiskers and do not meow. Moreover, for some correlated feature pairs, people have theories for why they are correlated, such as the fact that has wings is causally related to flies.

Thematic features provide a powerful means of activating feature relations and cuing theories about the object being categorized. For example, has wheels, manual transmission and drives in jungle clearly signifies some kind of vehicle and prompts any number of theories as to what kind. People in these kinds of experiments have enough information about the relations between features that generating theories about the category identity would be the natural tendency. In contrast, relations between attributes in the present experiment were much weaker in that features were very abstract and may not have correlated easily with one another and the category. For instance, unless the artificial label group clearly understood and/or knew the relationship between resonating frequency and volume they may have found it difficult to generate theories about the category label (indeed participants reported as much). Without clear theories, participants may have been less motivated to determine the category identity and instead may have memorized items and/or categorized based on similarity (e.g., dimensional values of items are similar to one another and their category). The picture emerging for the meaningful label group is somewhat different in that they already had a clear idea as to the category identity, based on the label. Moreover, the category label likely cues features most typical of the category. However, as the results suggest feature activation or connection may have been less likely to occur for the meaningful group. Perhaps category labels activated stored feature relations but abstract features interfered with these activations. Perhaps participants believing they already knew what features were predictive of the category, found it difficult to readjust their expectations to think of the category in new ways (see e.g., Keil, 2003, for a similar but slightly different observation). Thus, perhaps participants initially relied on prior knowledge when classifying, but found it difficult to map item information to prior expectations, and so, in the end, they chose to categorize based on memorization and similarity.

Other results were consistent with participants not only categorizing based on similarity but generalizing item information. Some recent research (Chin-Parker, & Ross, 2002; Chin-Parker & Ross, 2004; Markman & Ross, 2003; Yamauchi & Markman, 1998) may on some level explain participants' generalization tendencies. This research distinguishes between two types of category learning tasks, the standard category-learning task and the inference-based learning task. For the standard categorization task such as one used in this experiment, participants are given a full range of features and are asked to classify those features according to one of several categories. In this experiment participants were given lists of features and asked to categorize the list according to one of several types of instruments. In contrast, in the inference based learning task participants are presented with pictures (or lists) each missing a feature, along with the category label, and then asked to infer the missing feature. For example, a participant may be presented with a picture of a labeled bug (e.g., *Deezle*) and asked to predict a missing feature (e.g., *antenna*); the participant is then exposed to each of the features in this way until all the features are learned.

According to Markman and Ross (2003) classification learning leads to a focus on diagnostic features. Diagnostic features facilitate distinguishing one category from another. For instance, barking is particularly diagnostic of the category *dog* but not at all diagnostic of the category *cat*. Because classification learning focuses on specific features, representations formed are specific and exemplar like (Yamauchi & Markman, 1998). In contrast to classification learning, inference-learning demands learning entire feature lists within the context of the category label; this results in acquisition of both diagnostic and non-diagnostic features. Thus, you are not only going learn that dogs bark, but that cats do not bark, and that both have fur. As the emphases is on learning both diagnostic and non-diagnostic characteristics, performance on single-feature classification should be high and representations formed should be prototypical in kind.

Research by Chin-Parker and Ross (2004) provides a striking example of how inferring features leads to acquisition of general knowledge for a category (prototypical information) whereas classification learning promotes acquisition of diagnostic features (specific representations) and category differentiation. For their first experiment classification and inference learners learned two categories, one in which features diagnostic of category membership occurred eighty percent of the time, and one in which prototypical features overlapped. Test items included old items and transfer items that varied on distance to a category prototype. Performance on a forced choice test showed that participants in the classification condition were more likely to choose items with features highly diagnostic of category membership regardless of closeness to prototype, whereas inference learners showed no such preference. In their second experiment, classification and inference learners were exposed to similar categories; however instead of performing a forced choice test participants gave typicality ratings. Results revealed that the number of diagnostic features influenced judgments for classification learners whereas both diagnostic and non-diagnostic (but prototypical) features influenced judgments for inference learners. In sum results, showed that participants in the classification conditions learn specific category information, whereas inference learners tended to learn information that highlighted what the category is about.

These observations also extend to the present findings. For example, when categorizing at the superordinate level following transfer from basic levels participants failed to differentiate between prototype and fifty-fifty items (see also Appendix L). This would suggest that participants focused on diagnostic aspects of items without truly realizing the overall concept. Otherwise, they would have realized that fifty-fifty items did not belong to any one single category and categorization of these items would have been much more difficult. This expectation is even greater considering participants should have clearly seen that fifty-fifty items did not belong to anyone category when classifying at the basic level. Moreover, when summating false alarms (confusing basic level categories within a superordinate category with one another) and correct categorization, the prototype and fifty-fifty are almost identical even at the basic level. This further suggests that participants were focusing on diagnostic criteria and similarity of dimensional values to the category. They were simply focusing on the dimensional values most predictive of their category and failed to critically evaluate item information.

Meaningful labels were also expected to boost detection of taxonomic relations. Meaningful labels clearly defined taxonomic levels and were applied with the intent of making discovery of taxonomic relations easier. Learning the category *wind instrument* should have activated related subcategories, such as *flute* and *saxophone*. Clearly, the results failed to support a facilitation affect of meaningful labels. Indeed, in comparison to the artificial group reaction times were generally slower for the meaningful group especially when classifying at basic levels. One reason for this finding follows from the types of experiences cued by taxonomic labels and item features. When considering superordinate categories, both the label and attributes are on some level abstract. The category *wind instrument* does not necessarily specify any one instrument. Moreover, as noted previously dimensions are also abstract. When both the dimension and the category label are abstract prior experiences are in some respects unrestricted. That is, participants are free to consider or not consider specific instruments in which to attach item attributes. Indeed, all they really need to consider are general characteristics for *wind* and *percussion* instruments (note response times were at times marginally slower for the meaningful group so they may have considered specific features to some extent). Categorization experiences for the meaningful label group then are likely not all that different from those of the artificial group who make decisions without giving a great deal of consideration to specific categories. When considering experiences for the basic level a different picture emerges. Here the meaningful label group considered the specific category and the abstract item information was inconsistent with the categories. Hence, response times were much slower for the meaningful group when classifying at the basic level.

Finally, it is important to note that both the category label and generalization of item contribute to current findings. For instance, if the only factor-affecting participants' performance were prior knowledge, then performance for fifty-fifty items would be much poorer for the meaningful label group (following transfer to superordinate level from basic level). These items would be viewed as half-flute or half-saxophone and therefore odd. At the very least reaction times would be much slower. Conversely, if generalization were the only factor affecting results, then reaction time differences between artificial and meaningful label groups would be negligible. This is because the item structures were identical for both groups; thus prior knowledge cued by the category label is likely the primary factor differentiating performance. In short, several factors contribute to the current results none of which are independent.

7.3 Concluding remarks

The current research provides a good preliminary introduction into how abstract item structures interact with prior experience and taxonomic structures to influence category learning. However, additional research is needed to examine the parameters affecting this interaction. One implication of the present results is that people have deeply entrenched knowledge structures that can limit a label's extension to feature information, especially when that feature information contains little in the way of a thematic or correlated structure. Introducing thematic features is certainly one step toward further examining the relationship between prior experience and taxonomic category learning. For example, introducing the characteristic *brass* would likely promote an entirely different type of classification. If people tend to use taxonomic relationships then introducing thematic relations may facilitate this tendency.

One argument made several times in this paper is that categorization strategies used by the meaningful and artificial group may on some level be very similar in that participants in both groups relied on similarity judgments. Importantly, similarity judgments may be the preferred strategy for many learning situations. For instance, when learning new categories, contextual information is not always available. Under these circumstances determining how items are similar to one another and their category may be the only alternative.

These observations also extend to real world situations. In as much as stimuli used in this study extend to natural contexts, it is reasonable to assume that similarity judgments are a common and useful categorization strategy. As already noted the item structures used in this experiment are an extension of real world categories. Furthermore, similar to items used in this experiment it is common for real world objects to share attributes. Sometimes, the feature most common to certain breeds of *cats* is color. Pitch can be an important determinant for some types of *flutes*. Moreover, it is not unusual for domain specific objects of contrasting categories to have interchangeable features. Different breeds of *cats* often share the same color. *Flutes* and *flue pipes* are sometimes indistinguishable by pitch alone.

One consistent finding in this experiment is that when an object's identity was uncertain similarity judgments tended to be the rule rather than the exception. This may in part explain why conceptual preservation occurs. By conceptual preservation I mean that when learning new information it can be difficult to transcend prior experiences. There is a strong tendency, particularly when the concept is difficult, to attempt to incorporate new information into old conceptual structures. When new information is strongly connected to the prior experiences classification may be relatively easy. However, when the relationship is weak, the learning process may be interrupted or even terminated. These observations may explain why meaningful label participants in this study found it difficult to process basic level information. Perhaps they were attempting to classify from well-established prior experiences, but their experiences failed to extend clearly to new information. As a result, it may have been necessary for meaningful participants to rely on similarity judgments when classifying items.

Other results in this paper suggest that generalization may also be a key factor to early classification. Participants in both the meaningful and artificial label groups found classification easier when transferring from basic level conditions. As noted previously basic level categories that are in the same superordinate share a greater degree of similarity to one another than basic level categories that are in a different superordinate. Perhaps when participants learned basic level categories first they were in better position to learn structural characteristics of items and categories than were participants exposed to superordinate levels first. Participants learning basic categories were able to compare and contrast superordinate instantiations. They were able to compare how flute and saxophone were similar to one another and in turn how they were different from drum and bell. These comparisons may lead to not only a much deeper understanding of how dimensional values relate to each of the individual basic level categories but how they relate to superordinate categories. This is because having knowledge of how dimensional values connect to the basic level also translates into having knowledge how dimensional values connect to the superordinate level. Conversely, participants learning superordinate categories were only able to compare two categories. As a result they did not have foreknowledge of all basic level categories and had to consider how the dimensional values transferred to additional categories.

In sum, results of this study point to several trends. First, learning multiple categories increases the amount information compared sometimes making initial categorization more difficult, but may also lead to deeper understanding of how features transfer to new members. Second, when category information is ambiguous and difficult to attach to prior experiences similarity judgments may be the rule rather than the exception. Finally, prior experience may interact differently depending on the type of properties learned and the specificity of the taxonomic experience. When people have deeply embedded prior knowledge tied to basic level structures statistical properties may be more difficult to learn.

References

Ahn, W., Brewer, W.F., and Mooney, R.J. (1992). Schema acquisition form a single example. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 391-412.

Anderson, J.A. (1991). The adaptive nature of human categorization. Psychological Review, 98, 409-429.

Anglin, J.M. (1977). Word, Object, and Conceptual Development. New York: W. W. Norton.

Atran, S. (1990). Cognitive foundations of natural history. New York: Cambridge University Press.

Barsolou, L.W. (1983). Ad hoc categories. Memory and Cognition, 11, 211-227.

Barsolou, 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.

Barsolou, L.W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization. Cambridge: Cambridge University Press.

Barsolou, L.W. (1991). Deriving categories to achieve goals. In G.H. Bower (Ed.), The Psychology of Learning and Motivation, Vol.27 (pp 1 – 64). New York: Academic Press.

Brooks, L.W. (1992). Nonanalytic concept formation and memory for instance. In E. Rosch and B.B. Lloyd (Eds), Cognition and Categorization (pp. 169-211) New York: Wiley.

Bruner, J.S., Goodnow, J.J., and Austin G.A. (1956). A study of Thinking. New York: Wiley.

Carey, S. (1985). Conceptual Change in Childhood. Cambridge, M.A.: MIT Press.

Chomsky, N. (1957). Syntactic structures. The Hague. The Netherlands Mouton and Co.

Chin-Parker, S., and Ross, B.H. (2002). The effect of category learning on sensitivity to within-category correlations. Memory and Cognition, 30, 353-362.

Chin-Parker, S., and Ross, B.H. (2004). Diagnosticity and prototypicality in category learning: A comparison of inference learning and classification learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 30, 216-226.

Chomsky, N. (1957). Syntactic structures. The Hague, The Netherlands: Mouton and Co.

Coley, J.D. (1995). Emerging differentiation of folkbiology and folkpsychology. Child Development, 66, 1856-1874.

Coley, J.D., Medin, D.L., Proffit, J.B., Lynch, E., and Atran, S. (1999). Inductive reasoning inferences within folkbiological thought. In D.L. Medin and S. Atran (Eds), Folkbiology (pp. 205-232).

Collins, A.M., and Quillian, M.R. (1969). Retrieval time from semantic memory. Journal of Experimental Psychology: General, 127, 107-140.

Denton, M.D. (1985). Evolution: A Theory in Crises. Bethesda, MA: Adler and Adler Publishers.

Gelman, R. (2003). The Essential Child: Origins of Essentialism in Everyday Thought. New York: Oxford University Press.

Goodman, N. (1992/1972). Seven Structures on Similarity. In M. Douglas & D. Hull (Eds), How classification works, (pp. 13-23). Edinburgh, Scotland: Edinburgh University Press.

Gopnik , A., and Wellman, H.M. (1994). The Theory Theory. In L. Hirschfeld, and S. Gelman (Eds.), Mapping the mind: domain specificity in culture and cognition (pp. 257- 293). New York: Cambridge University Press.

Fried, L.S., and Holyoak, K.J. (1984). Induction of category distributions: A framework for classification learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 234-254.

Hampton, J.A. (1982). A demonstration of intransitivity in natural categories. Cognition, 12, 151-164.

Hampton, J.A. (1987). Inheritance of attributes in natural concept conjunctions. Memory and Cognition, 15, 55-71.

Hampton, J.A. (1995). Testing the prototype theory of concepts. Journal of Memory and Language, 34, 686-708.

Hampton, J. and Dubois, D. (1993). Psychological Models of Concepts: Introduction in Van Mechelen, et. al. (eds): Categories and Concepts: Theoretical views and inductive data analysis. Academic Press, London.

Heit, E. (1994). Models of the effects of prior knowledge on category learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 20, 1264-1282.

Heit, E., and Bott L. (2000). Knowledge selection in category learning. In K. Lamberts and D. L. Medin (Ed.), Psychology of Learning and Motivation (pp. 163-199). San Diego: Academic Press.

Heit, E., Briggs J., and Bott L. (2004). Modeling the effects of prior knowledge on learning incongruent features of category members. Journal of Experimental Psychology: Learning, Memory, and Cognition, 30, 1065-1081.

Hirschfeld, L.A., and Gelman, S.A. (Eds) (1994). Mapping the Mind: Domain Specificity in Cognition and Culture. Cambridge: Cambridge University Press.

Hull, C.L. (1920). Quantitative aspects of the evolution of concepts. Psychological Monographs, XXVIII.

Kaplan, A.S., and Murphy, G.L. (2000). Category learning with minimal prior knowledge. Journal of Experimental Psychology: Learning, Memory, and Cognition, 26, 829-846.

Keil, F.C., (1989). Concepts, Kinds, and Cognitive Development. Cambridge, MA: MIT Press.

Keil, F.C. (2003). Categorization, causation, and the limits of understanding. Language and Cognitive Processes, 18, 663-692.

Keil, F.C., and Rosenblit, L.R. (1997). Knowing Who Knows What. Paper presented at the 1997 meeting of Psychonomics Society, Philadelphia.

Komatsu, L.K. (1992). Recent views of conceptual structure. Psychological Bulletin, 112, 500-526.

Kruschke J.K. (2005). Category learning. In Lamberts, K. and Goldstone R.L, Handbook of Cognition. London: Sage Publication Inc.

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

Lutz, D.J., and Keil, F.C. (2002). Early understanding of the division of cognitive labor. Child Development, 73, 1073-1084.

Luria, A.R. (1970). The functional organization of the brain. Scientific American, 222, 66-78.

Malt, B.C. (1994). Water is not H_2O. Cognitive Psychology, 27, 41-70.

Malt, B.C. (1995). Category coherence in cross-cultural perspective. Cognitive Psychology, 29, 85-148.

Malt, B.C., and Johnson, E.C. (1992). Do artifact concepts have cores? Journal of Memory and Language, 31, 195-217.

Malt, B.C., and Smith, E.E. (1984). Correlated properties in nature categories. Journal of Verbal Learning and Verbal Behavior, 2, 250-269.

Markman, E.M. (1989). Categorization and naming in children. Cambridge, M.A: MIT Press.

Markman, E.M. (1990). Constraints children place in word meanings. Cognitive Science, 14, 57-77.

Markman, E.M., and Callanan, M.A. (1984). An analysis of hierarchical classification. In R. Sternberg (Ed.), Advances in the Psychology of Human Intelligence, Vol. 2 (pp. 325-365). Hillsdale, NJ: Erlbaum.

Markman, A.B., and Ross, B.H. (2003). Category use and category learning. Psychological Bulletin, 129, 592-613.

Medin, D.L., Altom, M.W., Edelson, S.M., and Freko, D. (1982). Correlated symptoms and simulated medical classification. Journal of Experimental Psychology: Learning, Memory, and Cognition, 8, 37-50.

Medin, D.L., and Atran, S. (2004). The native mind: Biological categorization and reasoning in development and across cultures. Psychological Review, 111, 960-983.

Medin, D.L., Lynch, E.B., and Salomon, K.O. (2000). Are there kinds of concepts? Annual Review Psychology, 51, 121-269.

Medin, D.L., Lynch, E.B., and Coley, J.D., and Atran, S., (1997). Categorization and reasoning among tree experts: Do all roads lead to Rome? Cognitive Psychology, 32, 49-96.

Medin, D.L., and Ortony, A., (1989). Psychological essentialism. In S. Vosniadou and A. Ortony (Eds.), Similarity and Analogical Reasoning (pp. 179-195). Cambridge: Cambridge University Press.

Medin, D.L., and Ross, B.H. (1989). The specific character of abstract thought: Categorization, problem solving, and induction. In R.J. Sternberg (Ed.), Advances in the psychology of human intelligence: Vol. 5, (pp. 189-223). Hillsdale, NJ: Erlbaum.

Medin, D.L., and Schaffer (1978). Context theory and classification learning. Psychological Review, 85, 307-338.

Medin, D.L., and Schwanenflugel (1981). Linear separability in classification learning. Journal of Experimental Psychology: Human Learning and Memory, 7, 355-368.

Medin, D.L, Wattenmaker, W.D., and Hampson, S.E. (1987). Family resemblance, conceptual cohesiveness, and category construction. Cognitive Psychology, 19, 242-279.

Mervis, C.B., and Rosch, E. (1981). Categorization of natural objects. Annual Review of Psychology, 32, 89-115.

Mervis, C.B., Catlin, J., and Rosch, E. (1976). Relationships among goodness-of-example, category, norms, and word frequency. Bulletin of the Psychonomic Society, 7, 283-284.

Murphy, G.L., (2002). The Big Book of Concepts. Cambridge, MA: MIT Press.

Murphy, G.L., and Allopenna, 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., and Andrew, J.M. (1993). The conceptual basis of antonymy and synonymy in adjectives. Journal of Memory and Language, 32, 301-313.

Murphy, G.L., and Brownell, H.H. (1985). Category differentiation in object recognition: Typicality constraints on the basic category advantage. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11, 70-84.

Murphy, G.L., and Kaplan, A.S. (2000). Feature distribution and background knowledge in category learning. Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 53A, 962-982.

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

Murphy, G.L., and Ross, B.H. (1999). Induction with cross-classified categories. Memory and cognition, 27, 1024- 1041.

Murphy, G.L., and Smith, E.E. (1982). Basic level superiority in picture categorization. Journal of Verbal Learning and Verbal Behavior, 21, 1-20.

Murphy, G.L., and Wisniewski E.J. (1989). Categorizing objects in isolation and in scenes: What a superordinate is good for. Journal of Experimental Psychology: Leaning, Memory, and Cognition, 15, 572-586.

Nakamura, G.V. (1985). Knowledge-based classification of ill-defined categories. Memory and Cognition, 13(5), 377-384.

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

Pazzani, M.J., (1991). Influence of prior knowledge on concept acquisition: Experimental and computational results. Journal of Experimental Psychology, 83, 353-363.

Posner, M.I., and Keele, S.W. (1968). On the genesis of abstract ideas. Journal of Experimental Psychology, 77, 353-363.

Posner, M.I., and Keele, S.W. (1970). On the genesis of abstract ideas. Journal of Experimental Psychology, 83, 304-308.

Proffitt, J.B., Coley, J.D., and Medin, D.L. (2000). Expertise and category-based induction. Journal of Experimental Psychology, 83, 367-379.

Osborne, G.J., and Calhoun, D.O. (1998). Themes, taxons, and trial types in children's matching to sample: Methodological considerations. Journal of Experimental Child Psychology, 68, 35-50.

Osherson, D.N., Smith, E.E., Wilkie, O., Lopez, A., and Shafir, E. (1990). Category-based induction. Psychological Review, 97, 185-200.

Quillian, M.R. (1967). Word concepts: A theory and simulation of some basic semantic capabilities. Behavioral Sciences, 12, 410-430.

Reed, S.K., (1972). Pattern recognition and categorization. Cognitive Psychology, 3, 382-407.

Rehder, B., and Murphy, G.L. (2003). A knowledge-resonance (KRES) model of category learning. Psychonomic Bulletin and Review, 10(4), 759-784.

Rips, L.J. (1989). Similarity, typicality, and categorization. In S. Vosniadou and A. Ortony (Eds.), Similarity and Analogical Reasoning (pp. 179-195). Cambridge: Cambridge University Press.

Rips, L.J., Shoben, E.J., and Smith, E.E. (1973). Semantic distance and the verification of semantic relations. Journal of Verbal Learning and Verbal Behavior, 12, 1-20.

Rosch, E. (1973a). Natural categories. Cognitive Psychology, 4, 328-350.

Rosch, E. (1973b). On the internal structure of perceptual and semantic categories. In T.E. Moore (Eds), Cognitive development and the acquisition of language. (pp. 111-144). New York: Academic Press.

Rosch, E. (1975). Cognitive representations of semantic categories. Journal of Experimental Psychology: General, 104, 192-233.

Rosch, E. (1978). Principles of categorization. In E. Rosch and B.B. Lloyd (Eds), Cognition and Categorization (pp. 27-48) Hillsdale, NJ: Erlbaum.

Rosch, E., and Mervis, C.B. (1975). Family resemblance: Studies in the internal structure of categories. Cognitive Psychology, 8, 573-605.

Rosch, E., and Mervis, C.B., Gray, W., Johnson, D., and Boyes-Braem, P. (1976). Basic objects in natural categories, Cognitive Psychology, *8*, 382-439.

Rosch, E., Simpson, C., and Miller, R.S. (1976). Structural bases of typicality effects. Journal of Experimental Psychology: Human Perception and Performance, 2, 491-502.

Rosenblit, L.R., and Keil, F.C. (2002). The misunderstood limits of folk science: An illusion of explanatory depth, Cognitive Science, 26, 521-562.

Ross, B.H. (1997). The use of categories affects classification. Journal of Memory and Language, 37, 240-267.

Ross, B.H., and Murphy, G.L. (1999). Food for thought: Cross-classification and category organization in a complex real-world domain. Cognitive Psychology, 38, 495-553.

Ross, B.H., and Spalding, T.L., (1994). Concepts and categories. In Thinking and Problem Solving: Academic Press.

Schyns, P.G., and Murphy, G.L. (1994). The ontogeny of part representation in object concepts. In D.L. Medin (Ed.), The Psychology of Learning and Motivation, Vol. 31 (pp. 305-349).

Sloman, S.A. (1993). Feature based induction. Cognitive Psychology, 25, 231-280.

Sloman, S.A. (1998). Categorical inference is no a tree: The myth of inheritance hierarchies. Cognitive Psychology, 35, 1-33.

Sloman, S.A., and Malt, B.C. (2003). Artifacts are not ascribed essences, nor are they treated as belonging to kinds. Language and Cognitive Processes,18, 563-582.

Smith, E.E., Balzano, G.J., and Walker, J. (1978). Nominal, perceptual, and semantic codes, in picture categorization. In J.W. Cotton and R.L. Klatzky (Eds.), Semantic Factors in Cognition (pp 137-168). Hillsdale, NJ: Erlbaum.

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

Smith, E.E., Osherson, D.N., Rips., and Keane, M. (1998). Combining prototypes: A selective modification model. Cognitive Science, 12, 485-527.

Smith, E.E., Shoben, E.J., and Rips, L.J. (1974). Structure and process in semantic memory: A selective model for semantic decisions. Psychological Review, 81, 214-241.

Smith, J.D., and Minda, J.P. (1998). Prototypes in the midst: The early epochs of category learning. Journal of Experimental of Psychology: Learning, Memory, and Cognition, 24, 1411-1436.

Smith, L. (1989). A model of perceptual classification in children and adults. Psychological Review, 96, 125-144.

Solomon, K.O., Medin, D.L., and Lynch, E. (1999. Concepts do more than categorize. Trends in Cognitive Sciences, 3, 99-105.

Spalding, T.L., and Murphy, G.L. (1996). Effects of background knowledge on category construction. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22, 525-538.

Storms, G., and De Boeck P. (1997). Formal models for intra-categorical structure that can be used for data analysis. In lamberts, Koen and Shanks, (Eds), Knowledge Concepts and Categories, Psychology Press.

Strevens, M. (2001). Only causation matters. Cognition, 82, 149-175.

Tanaka, J.W., and Taylor, M.E. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? Cognitive Psychology, 15, 121-149.

Wattenmaker, W.D. (1995). Knowledge structures and linear separability: Integrating information in object and social categorization. Cognition Psychology, 28, 274-328.

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

Waxman, S.R., Shipley, E.F., and Shepperson, B. (1991). Establishing new subcategories: The role of category labels. Child Development, 62, 127-138.

Wittgenstein, L. (1953). Philosophical Investigations. Oxford: Blackwell.

Wisniewski, E.J. (1995). Prior knowledge and functionally relevant features in concept learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21(2), 449-468.

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

Yamauchi, T., and Markman, A.B. (1998). Category learning by inference and classification. Journal of Memory and Language, 39, 124-148.

Appendix I

Untransformed means for the artificial group in Experiment 1

Taxonomic Group	Block 1	Block 2	Block3	Block4	Block 5	Block 6	Block 7	Block 8
Super-Super	.78	.81	.87	.89	.91	.90	.91	.92
Super-Basic	.76	.82	.85	.87	.46	.50	.56	.57
Basic-Super	.41	.51	.54	.56	.80	.83	.86	.90
Basic-Basic	.42	.48	.51	.57	.60	.62	.63	.67

Transformed means for the artificial group in Experiment 1

Taxonomic Group	Block 1	Block 2	Block3	Block4	Block 5	Block 6	Block 7	Block 8
Super-Super	.91	.99	1.12	1.17	1.20	1.20	1.22	1.24
Super-Basic	.87	.96	1.08	1.12	.48	.54	.60	.63
Basic-Super	.43	.54	.58	.61	.94	1.01	1.07	1.18
Basic-Basic	.44	.51	.54	.61	.65	.68	.69	.75

Appendix II

Untransformed means for the meaningful group in experiment 2

Taxonomic Group	Block 1	Block 2	Block3	Block4	Block 5	Block 6	Block 7	Block 8
Super-Super	.74	.83	.86	.87	.88	.89	.91	.93
Super-Basic	.76	.82	.83	.87	.45	.53	.57	.56
Basic-Super	.44	.52	.57	.59	.83	.87	.86	.90
Basic-Basic	.41	.49	.55	.57	.62	.65	.66	.67

Transformed means for the meaningful group in Experiment 2

Taxonomic Group	Block 1	Block 2	Block3	Block4	Block 5	Block 6	Block 7	Block 8
Super-Super	.84	.99	1.04	1.07	1.11	1.15	1.17	1.23
Super-Basic	.87	.97	1.02	1.08	.46	.56	.61	.59
Basic-Super	.45	.54	.59	.64	.97	1.07	1.09	1.16
Basic-Basic	.43	.52	.57	.62	.67	.70	.72	.77

Appendix III

Untransformed means for the artificial group in Experiment 3

Taxonomic Group	Block 1	Block 2	Block3	Block4	Block 5	Block 6	Block 7	Block 8
Super-Super	.74	.82	.85	.87	.90	.88	.90	.90
Super-Basic	.75	.82	.86	.88	.48	.54	.54	.58
Basic-Super	.44	.48	.53	.58	.82	.83	.86	.87
Basic-Basic	.45	.50	.55	.57	.61	.62	.65	.68

Untransformed means for the meaningful group in Experiment 3

Taxonomic Grouping	Block 1	Block 2	Block3	Block4	Block 5	Block 6	Block 7	Block 8
Super-Super	.75	.82	.84	.85	.86	.88	.91	.90
Super-Basic	.76	.82	.87	.88	.44	.46	.52	.56
Basic-Super	.41	.50	.49	.55	.84	.85	.87	.88
Basic-Basic	.40	.51	.58	.58	.60	.63	.65	.69

Transformed means for the artificial group in Experiment 3

Taxonomic Group	Block 1	Block 2	Block3	Block4	Block 5	Block 6	Block 7	Block 8
Super-Super	.86	1.00	1.08	1.13	1.18	1.18	1.19	1.18
Super-Basic	.86	.98	1.05	1.13	.50	.57	.58	.62
Basic-Super	.46	.51	.57	.64	1.02	1.03	1.08	1.10
Basic-Basic	.47	.52	.59	.62	.71	.70	.75	.77

Transformed means for the meaningful group in Experiment 3

Taxonomic Group	Block 1	Block 2	Block3	Block4	Block 5	Block 6	Block 7	Block 8
Super-Super	.86	.97	1.03	1.08	1.08	1.13	1.20	1.20
Super-Basic	.88	.98	1.08	1 .14	.46	.48	.58	.61
Basic-Super	.42	.52	.51	.59	1.02	1.05	1.10	1.12
Basic-Basic	.41	.53	.59	.65	.69	.73	.75	.78

Appendix IV

Untransformed means for the artificial group in Experiment 4

Taxonomic Group	Block 1	Block 2	Block3	Block4	Block 5	Block 6	Block 7	Block 8
Super-Super	.74	.80	.81	.83	.84	.85	.86	.87
Super-Basic	.75	.82	.84	.87	.47	.53	.55	.61
Basic-Super	.43	.50	.55	.59	.79	.86	.88	.87
Basic-Basic	.45	.50	.55	.62	.61	.64	.64	.67

Untransformed means for the meaningful group in Experiment 4

Taxonomic Group	Block 1	Block 2	Block3	Block4	Block 5	Block 6	Block 7	Block 8
Super-Super	.76	.80	.82	.87	.88	.89	.90	.90
Super-Basic	.75	.79	.80	.85	.47	.53	.57	.59
Basic-Super	.42	.48	.53	.57	.81	.81	.84	.87
Basic-Basic	.40	.50	.56	.58	.61	.62	.62	.65

Transformed means for the artificial group in Experiment 4

Taxonomic Group	Block 1	Block 2	Block3	Block4	Block 5	Block 6	Block 7	Block 8
Super-Super	.85	.94	.97	1.00	1.05	1.06	1.10	1.10
Super-Basic	.83	.93	.95	1.06	.49	.56	.59	.66
Basic-Super	.44	.52	.58	.64	.92	1.07	1.13	1.11
Basic-Basic	.46	.53	.58	.66	.67	.70	.73	.75

Transformed means for the meaningful group in Experiment 4

Taxonomic Group	Block 1	Block 2	Block3	Block4	Block 5	Block 6	Block 7	Block 8
Super-Super	.86	.94	.97	1.08	1.11	1.14	1.18	1.21
Super-Basic	.85	.91	.95	1.05	.49	.56	.61	.64
Basic-Super	.44	.50	.55	.58	.96	.99	1.03	1.09
Basic-Basic	.41	.51	.57	.60	.67	.69	.69	.71

Appendix V

5.1 Experiment 1 results for untransformed Data

5.1.1 Accuracy

For the following reasons arcsine transformations are employed in each of the five Experiments. Firstly, because proportions are bounded at zero at the low end of the scale and at one at the high end of the scale, they may not linearly relate to other continuous variables. Arcsine transforms dependent variables in the form of proportions by stretching out the tails of distributions of proportions. The arcsine transformation also has the added benefit of reducing violations of sphericity. For each of the 5 experiments two sets of analyses were performed for accuracy, one for untransformed data and for arcsine transformed data. Means for both untransformed and transformed data can be viewed in Appendix I-IV. Untransformed analyses can be viewed in Appendix V-IX. Arcsine transformed data are presented in the body of this paper.

Though arcsine transformations were employed it is also important to note several observations for the untransformed data. Proportions for untransformed accuracy data were all based on the same number of observations. Furthermore, the accuracies were reasonably normal, so restriction of range usually associated with accuracy proportions was not that much of a problem. Moreover, the variances were not as different as one might expect, however they at times differed enough to violate sphericity assumptions. Thus, while arcsine transformations were employed to help bring variances closer to an assumption of equality the actual violations of the sphericity were minimal.

Mixed factorial ANOVA's were performed on each set of analyses. The first set of analyses examines the question of perfect taxonomic transfer learning. The main question is whether participants who had no prior experience with categories can perform as well as participants who have had prior experience. For example, can participants learning basic-superordinate phase 2 categories perform as well as participants learning superordinate-superordinate phase 2 categories? A finding favoring this outcome would suggest a benefit to taxonomic learning as participants who have repeated categories (superordinate-superordinate phase 2 condition) have the clear advantage of seeing the same item and category structure over participants who have seen the same items but in the presence of a different category structure (basic-superordinate phase 2 condition).

Training block refers to learning over repeated blocks (e.g., within group or condition performance), group condition refers to comparisons made between taxonomic conditions (e.g., comparisons between basic-superordinate phase 2 condition and basic-basic phase 2 condition). Participants learned four training blocks in each condition.

First, focusing just on superordinate phase 2 conditions, analysis showed a significant main effect of training block ($\underline{F}$ (3, 132) = 10.22, $\underline{p}$ < .001) an interaction between training block and group ($\underline{F}$ (3, 132) = 8.23 $\underline{p}$ < .001) and a main effect of group condition ($\underline{F}$ (1, 44) = 9.01, $\underline{p}$ < .01). Participants in the superordinate-superordinate phase 2 condition performed better than participants in the basic-superordinate phase 2 condition on their respective first ($\underline{t}$ (44) = 20.42, $\underline{p}$ < .001) second, ($\underline{t}$ (44) = 8.56, $\underline{p}$ = .01) and third ($\underline{t}$ (44) = 3.34, $\underline{p}$ < .05) training blocks.

Next, focusing on basic phase 2 conditions, analysis showed a significant main effect of training block ($\underline{F}$ (3, 132) = 9.54, $\underline{p}$ < .001) and a main effect of group condition ($\underline{F}$ (3, 132) = 5.12 $\underline{p}$ < .001). Participants in the basic-basic phase 2 condition performed better than participants in the superordinate-basic phase 2 condition on their respective first ($\underline{t}$ (44) = 24.24 $\underline{p}$ < .001) second ($\underline{t}$ (44) = 7.35 $\underline{p}$ < .001) third ($\underline{t}$ (44) = 3.45, $\underline{p}$ < .05) and fourth ($\underline{t}$ (44) = 3.23, $\underline{p}$ < .05) training blocks.

Results failed to provide support for the first question asked in this section; participants transferring to new taxonomic levels (e.g., basic-superordinate phase 2 condition) did not perform as well as participants who repeated learning of taxonomic levels (e.g., superordinate-superordinate phase 2 condition).

The next set of analyses examines the question of whether any taxonomic learning occurred at all. Here all comparisons involve first time category exposures. The general idea is that if performance for participants learning phase 2 categories is superior to that of participants learning phase 1 categories there is evidence of taxonomic learning. That is, because participants in both conditions are learning particular categories for the first time, findings favoring taxonomic primed groups would suggest a learning advantage, due to experience with the taxonomically related category. This advantage may result from learning of class inclusion relations or any number of other factors. Given that participants are learning all categories for the first time comparisons are on some level standardized. However, it is important to note that participants learning the phase 2 categories have had prior exposure to items (i.e., the same items are presented in phase 1 and 2) thus may on some level have an advantage over phase 1 participants. Because there were no meaningful differences between superordinate-superordinate and superordinate-basic or between the basic-superordinate and basic-basic groups in phase 1, these pairs were combined for analysis. That is, the superodinate-basic phase 1 condition and the superodinate-superordinate phase 1 condition were combined as were the basic-superordinate phase 1 condition and the basic-basic phase 1 condition. Analyses therefore consisted of comparing the basic-superordinate phase 2 condition with the superordinate phase 1 combined condition and the superordinate-basic phase 2 condition with basic phase 1 combined condition.

First, focusing just on superordinate conditions, analysis showed a significant main effect of training block (F (3, 132) = 30.16, $p < .001$). No main effect of group condition was found; participants learning items for the first time in the basic-superordinate phase 2 condition failed (F (1, 44) = 1.01, $p > .12$) to outperform participants learning items for the first time in the superordinate phase 1 condition. The interaction between group condition and training block was also statistically non-significant (F (3, 132) = 1.40, $p > .53$).

Next, focusing on basic conditions, analysis showed a significant main effect of training block (F (3, 132) = 19.19, $p < .001$). No main effect of group condition was found, participants learning items for the first time in the superordinate-basic phase 2 condition failed (F (3, 132) = 1.21, $p > .08$) to outperform participants learning items for the first time in the basic phase 1 condition. The interaction between group condition and training block was also statistically non-significant (F (1, 44) = .54, $p > .40$)

Analyses failed to provide support for the second question asked in this experiment; participants learning categories for the first time in phase 2 did not outperform participants learning categories for the first time in phase 1. Thus, these comparisons did not support effects of taxonomic learning. In the next section reaction times are examined. Questions and expected outcomes for reaction times are identical to those for accuracy.

5.1.2 Reaction times.

The reaction times were averaged and submitted to mixed factorial ANOVA's, after discarding any times greater than 30 seconds. Only correct responses were analyzed. Analyses and comparisons are identical to those for accuracy. The first set of analyses examined the question of perfect taxonomic transfer by comparing first time category learning experiences with repeated category learning experiences. Effects of perfect taxonomic learning would show that participants learning categories for the first time in phase 2 perform as well as participants repeated the same categories in phase 2.

First, focusing on superordinate conditions, analysis showed a significant main effect of training block (F (3, 132) = 5.08, $p < .05$). The interaction between training block and group condition was statistically non significant (F (3, 132) = 1.96, $p > .09$) as was the main effect of group (F (1, 44) = .14, $p > .93$). No reaction time differences were evident between the basic-superordinate phase 2 condition and the superordinate-superordinate phase 2 condition.

Next, focusing on basic conditions, analysis showed a significant main effect of training block (F (3, 132) = 15.31, $p < .001$) and a main effect of group condition (F (1, 44) = 5.10, $p < .002$). Participants in the basic-basic phase 2 condition responded faster than participants in the superordinate-basic phase 2 condition on their respective first (t (44) = 3.51, $p < .001$) second (t (37) = 2.23, $p < .03$) third (t (44) = 2.26, $p < .03$) and fourth (t (44) = 2.52, $p < .02$).

Results showed partial support for effects of taxonomy in that participants who transferred from basic to superordinate categories processed items as fast as participants who transferred from superordinate to superordinate categories. Thus, participants repeating items in the presence of different taxonomic category performed as well as participants repeating items in presence of the same taxonomic category (indeed, though statistically non-significant they responded faster). However, this finding was apparent only for superordinate categories. Participants repeating basic level categories were faster than participants who transferred from superordinate to basic categories on all four blocks.

The next sets of analyses investigate whether there is any evidence of taxonomic learning. First, focusing just on superordinate conditions, analysis showed a significant main effect of training block (F (3, 132) = 7.82, $p < .001$), an interaction between training block and group condition (F (3, 132) = 5.10, $p < .002$) and a main effect of group condition (F (1, 44) = 10.11, $p < .003$). Participants learning items for the first time in the basic-superordinate phase 2 condition were faster processing items than participants learning items for the first time in the superordinate-basic phase 1 condition on their respective first (t (44) = 3.93, $p < .001$) second (t (44) = 2.17, $p < .04$) and third training blocks (t (44) = 2.52, $p < .02$).

Second, focusing just on basic conditions, analysis showed a significant main effect of training block (F (3, 132) = 7.82, $p < .001$) and an interaction between training block and group condition (F (3, 132) = 5.10, $p < .002$). Participants learning items for the first time in the superordinate-basic phase 2 condition were faster processing items than participants learning items for the first time in the basic-superodinate phase 1 condition on their respective first training block (t (44) = 2.24, $p < .03$).

When considering reaction times, the results generally favored the idea that learning categories for the first time following taxonomic transfer has advantages over learning items for the first time in phase 1, though it is important to note that processing advantages were greater for the basic-superordinate group than for the superordinate-basic group. The basic-superordinate group outperformed the superordinate-basic phase 1 condition on 3 blocks of training, whereas the superordinate-basic phase 2 condition outperformed the basic phase 1 condition on only the first block of training. As discussed next these differences in performance may follow in part from advantages to learning basic level categories first.

In summarizing differences between untransformed and transformed results for Experiment 1, no realized differences were found.

5.2 Experiment 2 results for untransformed data.

5.2.1 Accuracy

Mixed factorial ANOVA's were performed on each set of analyses. The first set of analyses examines the question of perfect taxonomic transfer learning. The main question here is whether meaningful labels will facilitate perfect transfer learning. If so then one would expect participants in the basic-superordinate phase 2 condition to perform as well as participants in the superordinate-superordinate phase 2 condition, and for participants in the superordinate-basic phase 2 condition to perform as well as participants in the basic-basic phase 2 condition.

First, focusing just on superordinate phase 2 conditions, analysis showed a significant main effect of training block ($\underline{F}$ (3, 126) = 10.95, $\underline{p} < .001$). The interaction between group condition and training block was statistically non-significant ($\underline{F}$ (3, 126) = .38, $\underline{p} > .80$), as was the main effect of group condition ($\underline{F}$ (1, 42) = 3.18 $\underline{p} < .08$).

Next, focusing on basic phase 2 conditions, analysis showed a significant main effect of training block (F (3, 132) = 11.74, $p < .001$) and a main effect of group condition (F (3, 132) = 6.12 $p < .001$). Participants in the basic-basic phase 2 condition performed better than participants in the superordinate-basic phase 2 condition on their respective first (t (43) = 32.74, $p < .001$) second (t (43) = 14.48, $p < .001$) third (t (43) = 8.12, $p < .01$) and fourth (t (43) = 14.36, $p < .001$) training blocks.

Results showed partial support for perfect taxonomic transfer effect in that participants in the basic-superordinate phase 2 condition performed nearly as well participants in superordinate-superordinate phase 2 condition. As discussed shortly this finding may reflect the generalization effect explored in Experiment 1.

The next set of analyses examines the question of whether any taxonomic learning occurred. Here all comparisons involve first time category exposures. If performance for participants learning phase 2 categories is superior to that of participants learning phase 1 categories there is evidence of taxonomic learning. Because there were no meaningful differences between superordinate-superordinate and superordinate-basic or between the basic-superordinate and basic-basic groups at phase 1, these pairs were combined for analysis.

First, focusing just on superordinate conditions, analysis showed a significant main effect of training block (F (3, 132) = 37.05, $p < .001$), and a main effect of group condition (F (1, 42) = 6.42, $p < .02$). Participants in the basic-superordinate phase 2 made a greater number of correct responses than participants in the superordinate phase 1 condition on their respective first (t (42) = 13.72, $p < .001$) and second (t (42) = 4.91, $p < .03$) training blocks.

Next, focusing on basic conditions, analysis showed a significant main effect of training block (F (3, 129) = 25.09, p < .001). The interaction between group condition and training block was statistically non-significant (F (3, 129) = 2.66, p = .06) as was the main effect of group condition (F (1, 43) = .03, p > .86). Participants learning items for the first time in the superordinate-basic phase 2 condition failed to outperform participants learning items for the first time in the basic phase 1 condition.

Analyses provided partial support for taxonomic transfer effect in that the basic-superordinate phase 2 condition outperformed the superordinate phase 1 condition. An effect of meaningful label may also be evident, as this finding was not found in Experiment 1.

5.2.2 Reaction times

The reaction times were averaged and submitted to mixed factorial ANOVA's, after discarding any times greater than 30 seconds. Only correct responses were analyzed. Analyses and comparisons are identical to those for accuracy. The first set of analyses examined the affect of meaningful labels on perfect taxonomic transfer.

First, focusing on superordinate conditions, analysis showed a significant main effect of training block (F (3, 126) = 4.43, p < .01). The interaction between training block and group condition was statistically non-significant (F (3, 126) = .51, p > .68) as was the main effect of group (F (1, 42) = 1.16, p > .07). Thus, no reaction time differences were evident between the basic-superordinate phase 2 condition and the superordinate-superordinate phase 2 condition.

Next, focusing on basic conditions, analysis showed a significant main effect of training block (F (3, 129) = 10.34, p < .001), as well as an interaction between group condition and training block (F (3, 129) = 6.77, $p < .001$) and a main effect of group condition (F (1, 42) = 8.55, $p < .01$). Participants in the basic-basic phase 2 condition responded faster than participants in the superordinate-basic phase 2 condition on their respective first (t (43) = 4.23, $p < .001$) and second (t (43) = 3.12, $p < .001$) training blocks.

Results showed partial support for perfect taxonomic transfer in that participants who transferred from basic to superordinate categories processed items as fast as participants who transferred from superordinate to superordinate categories. Thus, participants repeating items in the presence of different taxonomic category performed as well as participants repeating items in the presence of the same taxonomic category. However, this finding was apparent only for superordinate categories; participants repeating basic level categories were faster than participants who transferred from superordinate to basic categories on all four blocks. These findings replicate those found in Experiment 1.

The next analyses examine the possibility of any taxonomic transfer by comparing taxonomic transfer condition phase two performance to the performance in the matched conditions at phase 1. First, focusing just on superordinate conditions, analysis showed a significant main effect of training block (F (3, 132) = 8.21, $p < .001$) an interaction between training block and group condition (F (3, 132) = 5.43, $p < .001$) and a main effect of group condition (F (1, 44) = 9.23, $p < .003$). Participants learning items for the first time in the basic-superordinate phase 2 condition processed items faster than participants learning items for the first time in the superordinate phase 1 condition on their respective first (t (42) = 4.55, $p < .001$) second (t (42) = 2.64, $p < .03$) and third training blocks (t (42) = 2.15 $p < .04$).

Second, focusing just on basic conditions, analysis showed a significant main effect of training block (F (3, 129) = 37.12, $p < .001$). The interaction between training block and group condition was statistically non-significant (F (3, 129) = 2.00, $p > .21$) as was the main effect of group condition (F (1, 43) = .53, $p > .42$).

5.3 Experiment 3 results for untransformed means

5.3.1 Accuracy

The first set of analyses focuses on taxonomic transfer effects. Examination of taxonomic transfer effects in the present experiment focuses solely on comparisons between the artificial and meaningful groups in the basic-superodinate phase 2 conditions and superordinate-basic phase 2 conditions. These comparisons show whether one group benefits more than the other from taxonomic priming. A finding favoring the meaningful group would also suggest that instructions boosted the manipulation of knowledge associated with the meaningful label.

First when comparing mean transfer differences between the meaningful basic-superordinate phase 2 condition and the artificial basic superordinate phase 2 condition results showed a significant main effect of training block (F (3, 243) = 29.34, $p < .001$). The interaction between group condition and training block was statistically non-significant (F (3, 243) = .13, $p > .43$) as was the main effect of group (F (1, 81) = 1.77, $p > .17$).

Next, when comparing mean transfer between meaningful superordinate-basic phase 2 condition and the artificial superordinate-basic phase 2 condition results showed a significant main effect of training block (F (3, 243) = 53.60, $p < .001$). The interaction between group condition and training block was statistically non-significant (F (3, 243) = .20, $p > .92$) as was the main effect of group (F (1, 81) = 2.96, $p > .10$).

In sum, no differences were found between meaningful and artificial label groups following transfer to a new taxonomic level.

The next set of comparisons examine phase 1 and phase 2 differences between artificial and meaningful groups. These comparisons include superordinate and basic level phase 1 conditions, as well as superordinate and basic level phase 2 conditions. These comparisons are important for further examining whether instructions facilitate learning for the meaningful label group. Findings showing better performance for the meaningful group would support this idea.

First, focusing on superordinate phase 1 conditions for artificial and meaningful groups, results showed a significant main effect of training block (F (3, 117) = 55.73, $p < .001$). The interaction between group condition and training block was statistically non-significant (F (3, 117) = 1.96, $p > .14$) as was the main effect of group (F (1, 39) = 1.40, $p > .24$).

Second, focusing on basic phase 1 conditions for artificial and meaningful groups, results showed a significant main effect of training block (F (3, 117) = 122.33, $p < .001$). The interaction between group condition and training block was statistically non-significant (F (3, 117) = 1.75, $p > .10$) as was the main effect of group (F (1, 39) = 1.55, $p > .14$).

Third, comparisons between the artificial superordinate-superordinate phase 2 condition with the meaningful superordinate-superordinate phase 2 conditions showed a significant main effect of training block ($\underline{F}$ (3, 117) = 6.44, $\underline{p}$ < .02). The interaction between group condition and training block was statistically non-significant ($\underline{F}$ (3, 117) = .88, $\underline{p}$ > .29) as was the main effect of group ($\underline{F}$ (1, 39) = 2.76, $\underline{p}$ > .19).

Finally, when comparing the artificial basic-basic phase 2 condition with the meaningful basic-basic phase 2 condition, results showed a significant main effect of training block ($\underline{F}$ (3, 117) = 5.90, $\underline{p}$ < 001). The interaction between group condition and training block was statistically non-significant ($\underline{F}$ (3, 117) = 1.21, $\underline{p}$ > .24) as was the main effect of group condition ($\underline{F}$ (1, 39) = .88, $\underline{p}$ > .35).

In sum, analyses revealed no differences between artificial and meaningful groups for correct responses. Instructions did not have the expected effect of boosting knowledge effects associated with the meaningful labels.

5.3.2 Reaction time*s*

The reaction times were averaged and submitted to mixed factorial ANOVA's, after discarding any times greater than 3 seconds. Only correct responses were analyzed. It is important to keep in mind that participants learn in total 48 items for each block, 12 for each category.

The first sets of comparisons examine taxonomic transfer differences between the artificial and meaningful label groups for basic-superordinate phase 2 conditions and superordinate-basic phase 2 conditions. Expectations for these comparisons are identical to those for correct responses.

First, when comparing the meaningful basic-superordinate phase 2 and artificial basic superordinate phase 2 condition results showed a significant main effect of training block ($\underline{F}$ (3, 243) = 24.43, $\underline{p}$ < .001). The interaction between group condition and training block was statistically significant (F (3, 243) = 3.65, p < .01). Participants in the artificial basic-superordinate condition were faster processing items at block 1 (t (81) = 3.96, p = 6.11, p < .02).

Next, comparisons between the meaningful superordinate-basic phase 2 condition and the artificial superordinate-basic phase 2 condition showed a significant main effect of training block (F (3, 243) = 21.22, p < .001) an interaction between group and training block (F (3, 243) = 10.44, p < .001) as well as a main effect of group (F (1, 81) = 20.74, p < .001). Participants in the artificial superordinate-basic phase 2 condition were faster processing standard items on their respective first (t (81) = 26.55, p < .001) second, (t (81) = 12.13, p < .001) third, (t (81) = 12.91, p < .001), and fourth (t (81) = 10.65, p < .003) training blocks.

In sum, the artificial group was faster than the meaningful when transferring from one taxonomic level to another regardless of type of structure first learned. Clearly instructions failed to boost knowledge effects associated with the meaningful label.

The next set of comparisons examine phase 1 and phase 2 differences between artificial and meaningful groups. These comparisons include superordinate and basic level phase 1 conditions, as well as superordinate and basic level phase 2 conditions.

First, focusing on the artificial superordinate-superordinate phase 1 condition with the meaningful superordinate-superordinate phase 1 condition results showed a significant main effect of training block (F (3, 117) = 39.24, p < .001) and an interaction between training block and group condition (F (3, 117) = 3.73, p < .02). The main effect of group was statistically non-significant (F (1, 39) = 2.01, p > .10). Participants in the artificial label group processed items faster on training block 1 (t (39) = 2.93, p < .03).

Second, when comparing the artificial basic-basic phase 1 condition with the meaningful basic-basic label phase 1 condition, results showed a significant main effect of training block (F (3, 117) = 27.11, $p < .001$) an interaction between group and training block (F (3, 117) = 10.81, $p < .001$) as well as a main effect of group (F (1, 39) = 32.11, $p < .001$). Participants in artificial basic-basic group were faster than participants in the meaningful basic-basic group when processing items on training blocks 1 (t (39) = 6.93, $p < .001$), 2 (t (39) = 5.25, $p < .001$), 3 (t (39) = 4.50, $p < .001$) and 4 (t (39) = 3.90, $p < .001$).

Third, when comparing the artificial superordinate-superordinate phase 2 condition with the meaningful superordinate-superordinate phase 2 condition results showed a significant main effect of training block (F (3, 117) = 3.26, $p < .03$). The interaction between training block and group condition was statistically non-significant, (F (3, 117) = .43 $p > .73$) as was the main effect of group condition (F (1, 39) = .01, $p > .93$).

Finally, focusing on the artificial basic-basic phase 2 condition and the meaningful basic-basic label phase 2 condition results showed a significant main effect of training block (F (3, 117) = 15.49, $p < .001$). The interaction between training block and group condition was statistically non-significant (F (3, 117) = 1.09, $p > .35$) as was the main effect of group condition (F (1, 39) = 1.21, $p > .31$).

5.4 Experiment 4 results

5.4.1 Artificial and meaningful groups comparisons

The following sets of comparisons are important for verifying results found in the Experiment 3. In that Experiment no differences were found between artificial and meaningful groups on measures of accuracy. However, reaction time performance was overall much faster for the artificial label group. One factor that may have contributed to weak performance on the part of the meaningful group, particularly with respect to reaction time performance was the inclusion of instructions in that Experiment. Participants in the meaningful group having knowledge of categories may have spent more time trying to figure out how one feature was related to the other. Replicating the analysis performed in Experiment 3 is important for ruling out this possibility. The first sets of comparisons focus on learning of standard items over the first eight blocks of training. Block 9 comparisons for standard, prototype, and fifty-fifty items are explored later.

5.4.1.1 Accuracy

The first set of comparisons examine phase 1 and phase 2 differences between artificial and meaningful groups. These comparisons include superordinate and basic level phase 1 conditions, as well as superordinate and basic level phase 2 conditions. Findings favoring the meaningful group would show that meaning attached to the label boost classification for that group.

First, when comparing the artificial superordinate-superordinate phase 1 condition with the meaningful superordinate-superordinate phase 1 condition results showed a significant main effect of training block ($\underline{F}$ (3, 150) = 60.22, $\underline{p}$ < .001). The interaction between group and training block was statistically non-significant ($\underline{F}$ (3, 150) = 1.88, $\underline{p}$ > .44) as was the main effect of group ($\underline{F}$ (1, 50) = 1.60, $\underline{p}$ > .18).

Next, when comparing the artificial basic-basic label phase 1 condition with the basic-basic meaningful label phase 1 condition results showed a significant main effect of training block ($\underline{F}$ (3, 150) = 134.77, $\underline{p}$ < .001). The interaction between training block and group was statistically non-significant ($\underline{F}$ (3, 150) = 1.29, $\underline{p}$ > .12) as was the main effect of group ($\underline{F}$ (1, 50) = 1.66, $\underline{p}$ > .15).

Third, when comparing the artificial superordinate-superordinate phase 2 condition with the meaningful superordinate-superordinate phase 2 conditions, results, showed a significant main effect of training block ($\underline{F}$ (3, 150) = 6.78, $\underline{p}$ < .01). The interaction between group and training block was statistically non-significant ($\underline{F}$ (3, 150) = .56, $\underline{p}$ > .65) as was the main effect of group ($\underline{F}$ (1, 50) = 2.24, $\underline{p}$ > .14).

Finally, focusing on the artificial basic-basic phase 2 condition and the meaningful basic-basic phase 2 condition, results showed a significant main effect of training block ($\underline{F}$ (3, 150) = 5.22, $\underline{p}$ < 003). The interaction between group condition and training block was statistically non-significant ($\underline{F}$ (3, 150) = 1.02, $\underline{p}$ > .33), as was the main effect of group condition ($\underline{F}$ (1, 50) = .17, $\underline{p}$ > .49).

5.4.1.2 Reaction times

Possible outcomes for group reaction times are several. Faster performance on part of the meaningful group would not only suggest that meaning attached to the label boosted performance for that group, but that instructions presented in Experiment 3 adversely affected performance of the meaningful group. Conversely, a replication of findings in Experiment 3, that is slower reaction times for the meaningful label, would support the idea that meaning attached to label interacts adversely with abstract dimensions and incongruent dimensional values.

The next set of comparisons examine phase 1 and phase 2 differences between artificial and meaningful groups. These comparisons include superordinate and basic level phase 1 conditions, as well as superordinate and basic level phase 2 conditions.

First, when comparing the artificial superordinate-superordinate phase 1 condition with the meaningful superordinate-superordinate phase 1 condition, results showed a significant main effect of training block (F (3, 150) = 45.91, $p < .001$). The interaction between training block and group condition was statistically non-significant (F (3, 150) = 3.55, $p > .05$) as was the main effect of group condition (F (1, 150) = 1.78, $p > .12$).

Second, focusing on the artificial basic-basic phase 1 condition and the meaningful basic-basic phase 1, results showed significant main effect of training block, F (3, 150) = 51.28, $p < .001$, a significant interaction between training block and group condition (F (3, 150) = 10.22, $p < .001$) as well as a main effect of group condition (F (1, 50) = 23.77, $p < .001$). The artificial basic phase 1 condition was significantly faster processing items on blocks one (t (50) = 5.85, $p < .001$), two (t (50) = 3.91, $p < .001$), three (t (50) = 3.55, $p < .001$, and four (t (50) = 3.01, $p < .01$).

Third, focusing on the artificial superordinate-superordinate phase 1 condition with the meaningful superordinate-superordinate phase 1 condition showed a significant main effect of training block (F (3, 150) = 8.21, $p < .001$). The interaction between training block and group condition was statistically non-significant (F (3, 150) = 3.01, $p > .07$) as was the main effect of group condition (F (1, 50) = 1.01, $p > .44$).

Finally, when comparing the artificial basic-basic phase 2 condition with the meaningful basic-basic phase 2 condition results showed a significant main effect of training block (F (3, 150) = 10.22, $p < .001$). The interaction between training block and group condition was statistically non-significant (F (3, 350) = 2.90, $p > .06$), as was the main effect of group condition (F (1, 50) = 1.77, $p > .09$).

In sum, findings replicate those of Experiment 3. Participants in the meaningful group performed reliably slower than participants in the artificial group, but only when learning the basic-basic phase 1 condition. This finding suggests that meaning attached to the meaningful label interacts adversely with prior expectations of participants.

5.4.2 Artificial and meaningful label group comparisons for block 9

The following analyses compare group differences when learning standard, prototype, and fifty-fifty items at block 9. As prototype items are congruent with participants prior expectations one would expect enhanced performance on part of the meaningful group for these items. However, if abstract dimensions interact adversely with prior experiences, weaker performance (or equal performance) is expected on part of the meaningful group for these items.

5.4.2.1 Accuracy

First, when comparing the artificial superordinate-superordinate phase 2 condition with the meaningful superordinate-superordinate phase 2 condition, results showed a significant main effect of item type, (F (2, 100) = 56.59, $p < .001$). The interaction between group and item type was statistically non-significant (F (2, 100) = 1.31, $p > .25$), as was the main effect of group (F (1, 50) = 1.56, $p > .20$).

Second, focusing on the artificial basic-superordinate phase 2 condition and the meaningful basic-superordinate phase 2 condition, results, showed a significant main effect of item type (F (2, 100) = 49.35, $p < .001$). The interaction between group and item type was statistically non-significant (F (2, 100) = .17, p > .90) as was the main effect of group (F (1, 50) = .25, p > .81).

Third, when comparing the artificial basic-basic label phase 1 condition with the basic-basic meaningful label phase 1, results showed a significant main effect of item type (F (2, 100) = 59.53, p < .001). The interaction between training block and group was statistically non-significant (F (2, 100) = .70, p > .44) as was the main effect of group (F (1, 50) = 1.11, p > .20).

Finally, focusing on the artificial basic-basic phase 2 condition and the meaningful basic-basic phase 2 condition, results showed a significant main effect of training block (F (2, 100) = 32.33, p < .001). The interaction between group condition and training block was statistically non-significant (F (2, 100) = .94, p > .39) as was the main effect of group condition (F (1, 50) = .50, p > .40).

In sum, no differences were found between groups when learning prototype, fifty-fifty, and standard items. This would suggest that even when features are congruent with prior experience as with prototype items, the meaningful group fails to benefit.

5.4.2.2 Reaction times

First, focusing on the artificial superordinate-superordinate phase 2 condition and the meaningful superordinate-superordinate phase 2 condition results showed a significant main effect of item type, F (2, 100) = 4.52, p < .001. The interaction between group condition and item type was statistically non-significant, F (2, 100) = .25, p > .83, as was the main effect of group condition, F (1, 50) = .34, p > .23.

Second, when comparing the artificial basic-superordinate phase 2 condition and the meaningful basic-superordinate phase 2 condition, results showed a significant main effect of item type ($\underline{F}$ (2, 100) = 19.42, $\underline{p}$ < .001). The interaction between group condition and item type was statistically non-significant ($\underline{F}$ (2, 100) = 3.19, $\underline{p}$ > .05) as was the main effect of group condition ($\underline{F}$ (1, 50) = 2.89, $\underline{p}$ > .10).

Third, when comparing the basic-basic phase 2 condition with the meaningful basic-basic phase 2 condition, results showed a significant main effect of item type ($\underline{F}$ (2, 100) = 4.67, $\underline{p}$ < .02). The interaction between group and item type was statistically non-significant ($\underline{F}$ (2, 100) = .1.09, $\underline{p}$ > .25) as was the main effect of group condition ($\underline{F}$ (1, 50) = 1.90, $\underline{p}$ > .14). Note, that an independent sample t-test showed that the artificial label group was faster processing standard items ($\underline{t}$ (50) = 3.00, $\underline{p}$ < .03).

Finally, when focusing on the artificial superordinate-basic phase 2 condition the meaningful superordinate-basic phase 2 condition results determined that the main effect of item type was statistically non-significant ($\underline{F}$ (2, 100) = 2.91, $\underline{p}$ > .05) as was the interaction between group condition and item type ($\underline{F}$ (2, 100) = 1.87, $\underline{p}$ > .10) and the main effect of group condition ($\underline{F}$ (1, 50) = 1.70, $\underline{p}$ > .09).

5.4.3 Block 9 item comparisons for individual groups

5.4.3.1 Artificial label.

The following sets of analyses examine differences between items for block 9. These analyses are important for exploring how differences in item structure affect learning of items. Participants are expected to prefer the structural qualities of prototype items as compared to standard items. Expectations for prototype and fifty-fifty items are less clear. If participants categorized based on the number of dimensional values that belong to superordinate categories, then small differences are expected for these items. On the other hand, if they classify based on prior knowledge poorer performance is expected for fifty-fifty items. Comparisons are made first for the artificial label and then for the meaningful label.

5.4.3.2 Accuracy

Prototype items were classified better than standard items in both the superordinate-superordinate phase 2 ($\underline{t}$ (25) = 2.44, $\underline{p} < .001$) and basic-superordinate phase 2 conditions ($\underline{t}$ (25) = 7.77, $\underline{p} < .001$). Prototype items were also classified better than standard items in both the basic-basic phase 2 ($\underline{t}$ (25) = 9.79, $\underline{p} < .001$) and superordinate-basic phase 2 conditions ($\underline{t}$ (25) = 9.00, $\underline{p} < .001$).

No differences were found between prototype and fifty-fifty items in either the superordinate-superordinate phase 2 ($\underline{t}$ (25) = 1.01, $\underline{p} = .30$) and the basic-superordinate phase 2 conditions ($\underline{t}$ (25) = 1.34, $\underline{p} > 18$).

Fifty-fifty items were classified better than standard items in both the superordinate-superordinate phase 2 ($\underline{t}$ (25) = 5.00, $\underline{p} < .001$) and the basic-superordinate phase 2 conditons ($\underline{t}$ (25) = 6.79, $\underline{p} < .001$).

5.4.3.3 Reaction times for item comparisons

Prototype items were classified faster than standard items when classifying in the superordinate-superordinate phase 2 level ($\underline{t}$ (25) = 2.78, $\underline{p} < .01$). However, no reaction time differences were found between these items when classifying in the basic-superordinate phase 2 condition ($\underline{t}$ (25) = .88, $\underline{p} = .72$). No reaction time differences were found between prototypes and standard items in either the basic-basic phase 2 condition ($\underline{t}$ (25) = 2.05, $\underline{p} > .05$) or the superordinate-basic phase 2 condition ($\underline{t}$ (25) = .89, $\underline{p} > .30$). No reaction time differences were found between prototype and fifty-fifty items in either the superordinate-superordinate phase 2 condition, ($\underline{t}$ (25) = 1.50, $\underline{p} > .06$), or basic-superordinate phase 2 condition (t (25) = .98, p > .37). No reaction time differences were found between standard and fifty-fifty items in either the superordinate-superordinate phase 2 condition (t (25) = 1.00, p > .23) or the basic-superordinate phase 2 condition (t (25) = 1.70, p > .10).

5.4.4 Meaningful Label

Meaningful group participants are expected to classify prototype items better than both standard items. Moreover, in comparison to standard items and prototype items poorer performance is expected for fifty-fifty items. This is because the combination of dimensional values for fifty-fifty is inconsistent with participant's prior expectations for instruments.

5.4.4.1 Accuracy

Prototype items were classified better standard items in both the superordinate-superordinate phase 2 condition (t (25) = 5.24, p < .001) and basic-superordinate phase 2 condition (t (25) = 8.77, p < .001). Prototype items were classified better than standard items in both the basic-basic phase condition (t (25) = 8.98, p < .001) and superordinate-basic phase 2 condition, (t (25) = 7.73, p < .001). No differences were found between prototype and fifty-fifty items when classifying these items in the superordinate-superordinate phase 2 condition (t (25) = 1.81, p > .08) or the basic-superordinate phase 2 condition (t (25) = .27, p > .79). Fifty-fifty items were classified better than standard items in both the superordinate-superordinate phase 2 condition (t (25) = 7.99, p < .001) and the basic-superordinate phase 2 condition (t (25) = 7.55, p < .001).

When considering accuracy findings at block 9, performances for prototype and fifty-fifty items was generally better than for standard items. Moreover, no differences were found between prototype and fifty-fifty items. The findings are consistent with the idea that participants are making classification decisions based on the number of dimensional values that correctly predict the category. Thus participants do not appear to be making decisions based on prior experiences with instruments (otherwise a half flute/half saxophone would seem odd in comparison to an instrument that is all flute or mostly flute).

5.4.4.2 Reaction times

Prototype items were categorized faster than standard items in the superordinate-superordinate phase 2 condition (t (25) = 4.22, $p < .001$). However, no differences were found between these items when categorizing in the basic-superordinate phase 2 condition (t (23) = 1.45, $p > .50$). Prototype items were classified faster than the standard items at both the basic-basic phase 2 condition (t (25) = 6.45, $p > .001$) and the basic-superordinate phase 2 condition (t (25) = 3.33, $p < .001$). No reaction time differences were found between prototype and fifty-fifty items when categorizing in either the superordinate-superordinate phase 2 condition (t (25) = 2.00, $p > .05$) and the superordinate-basic phase 2 condition (t (25) = 1.89, $p > .10$). Fifty-fifty items were categorized faster than standard items in the basic-superordinate phase 2 condition (t (23) = 2.98, $p < .01$). However no reaction time differences were found between these *items in* the superordinate-superordinate phase 2 condition (t (25) = 1.55, $p > .20$).

5.4.5 Differences between taxonomic levels

Then next sets of analyses examine mean differences between superordinate and basic level categories for artificial and meaningful groups. These analyses are important for differentiating the influence of prior expectations, dimensions, and item structure, on categorization. For example, slower responses on part of the meaningful group for prototype items would suggest that abstract dimensions negatively impacted performance. This is because the primary factor affecting performance for prototype items are abstract dimensions (structure for prototype items was held constant between taxonomic levels). However, slower responses to standard items on part of the meaningful group would suggest that both abstract dimensions and incongruent dimensional values negatively affected meaningful group performance (these items have both abstract dimensions and incongruent dimensional values). Finally, slower between taxonomic level responses to standard than to prototype items on would suggest that incongruent dimensional values are the primary factor affecting performance. The first set of analyses compares mean differences between basic-basic and superordinate-superordinate groups. The second set of analyses examines group mean differences between basic-superordinate and superordinate-basic phase 2 conditions.

First, when comparing differences between basic-basic and superordinate-superordinate groups for standard items results showed mean differences were smaller for the artificial label on blocks, 1, 2, 3, and 5 ($p < .01$). Comparisons for block 9 failed to find mean reaction time differences between groups for standard ($t\ (50) = 1.98$, $p > .15$) prototype ($t\ (50) = 1.05$, $p > .23$). Moreover, no between taxonomic level differences were found when comparing prototype and standard items ($t\ (50) = 1.02$, $p > .14$).

Next, when comparing differences between basic-superordinate and superordinate-basic phase 2 condition for standard items no mean differences were found between groups on block, 5, 6, 7, and 8 ($p > .77$). Moreover comparisons for block 9 also failed to reveal mean reaction time differences between groups for standard ($t\ (50) =$

1.57, $p > .05$) and prototype items (t (50) = 1.78, $p > .06$). No between taxonomic level differences were found when comparing prototype and standard items (t (50) = .72, $p > .53$).

In summarizing differences between untransformed and transformed groups, no realized differences were found.

5.5 Experiment 5 results

5.5.1 Item comparisons for artificial and meaningful groups

The following analyses compare group differences when learning standard, prototype, and fifty-fifty items at blocks 1, 5, and 9. As prototype items are congruent with participants prior expectations one would expect enhanced performance on part of the meaningful group for these items. However, if abstract dimensions interact adversely with prior experiences, weaker performance (or equal performance) is expected on part of the meaningful group for these items.

5.5.1.1 Accuracy

The first sets of analyses compare artificial and meaningful label groups for prototype, fifty-fifty, and standard items, on blocks 1, 5, and 9 for superordinate level categories.

First, when comparing the artificial superordinate-superordinate block 1 condition with the meaningful superordinate-superordinate block 1 condition, results showed a main effect of item type (F (2, 108) = 77.77, $p < .001$). However, the interaction between item type and group was statistically non-significant (F (2, 108) = 1.05, $p > .18$), as was the main effect of group (F (1, 54) = .97, $p > .35$).

Second, when comparing the artificial superordinate-superordinate block 5 condition with the meaningful superordinate-superordinate block 5 condition, results showed a main effect of item type (F (2, 108) = 55.45, p < .001). The interaction between item type and group was statistically non-significant (F (2, 108) = .30, p > .76) as was the main effect of group (F (1, 54) = 2.00, p > .20).

Finally, when comparing the artificial superordinate-superordinate block 9 condition with the meaningful superordinate-superordinate block 9 condition, results showed a main effect of item type (F (2, 108) = 60.32, p < .001). The interaction between item type and group was statistically non-significant (F (2, 108) = .23, p > .66), as was the main effect of group (F (1, 54) = 2.00, p > .22).

In sum, the artificial and meaningful labels did not differ in performance on prototype, standard, and fifty-fifty items at block 1, 5, and 9, for the superordinate-superordinate phase 1 and 2 conditions.

The next sets of analyses compare artificial and meaningful label groups for prototype, fifty-fifty, and standard items, on blocks 1, 5, and 9 for basic level categories.

First, when comparing the artificial basic-basic group with the meaningful basic-basic group for block 1, results showed a main effect of item type (F (2, 108) = 20.71, p < .001). The interaction between item type and group was statistically non-significant (F (2, 108) = 1.00, p > .44) as was the main effect of group (F (1, 54) = .56, p > .56).

Second, when comparing the artificial basic-basic group with the meaningful basic-basic group for block 5, results showed a main effect of item type (F (2, 108) = 25.20, p < .001). The interaction between item type and group was statistically non-significant (F (2, 108) = .09, p > .95), as was the main effect of group (F (1, 54) = .05, p > .77).

Finally, when comparing the artificial basic-basic group with the meaningful basic-basic group for block 9, results showed a main effect of item type (F (2, 108) = 77.33, p < .001). The interaction between item type and group was statistically non-significant (F (2, 108) = .88, p > .20) as was the main effect of group (F (1, 54) = .02, p > .95).

In sum, the artificial and meaningful labels did not differ in performance on prototype, standard, and fifty-fifty items at block 1, 5, and 9, for the basic-basic phase 1 and 2 conditions

The next sets of comparisons explore differences between groups for taxonomic transfer phase 2 categories.

First, focusing on basic-superordinate phase 2 condition for block 5 results showed a main effect of item type (F (2, 108) = 100.73, p < .001. The interaction between item type and group was statistically non-significant, F (2, 108) = .14, p > .97, as was the main effect of group (F (1, 54) = .44, p = .55). Second, focusing on block 9, results showed a main effect of item type (F (2, 108) = 77.09, p < .001). The interaction between item type and group was statistically non-significant (F (2, 108) = .98, p > .67) as was the main effect of group (F (1, 54) = 1.01, p > .39).

Next, focusing on superordinate-basic phase 2 condition for block 5, results showed a main effect of item type (F (2, 108) = 13.60, p < .001). The interaction between item type and group was statistically non-significant (F (2, 108) = .55, p > .74) as was the main effect of group (F (1, 54) = .56, p > .77). Second, focusing on block 9, results showed a main effect of item type (F (2, 108) = 62.99, $p < .001$). The interaction between item type and group was statistically non-significant (F (2, 108) = .07, $p > .97$) as was the main effect of group (F (1, 54) = .05, $p > .87$).

In sum, results failed to reveal reliable differences between groups when categorizing items at either superordinate or basic levels.

5.5.1.2 Reaction Times

The first sets of analyses compare artificial and meaningful label groups for prototype, fifty-fifty, and standard items, on blocks 1, 5, and 9, for superordinate level categories.

First, when comparing the artificial superordinate-superordinate group with the meaningful superordinate-superordinate group at block 1, results showed a main effect of item type (F (2, 108) = 6.78, $p < .01$). The interaction between item type and group was statistically non-significant (F (2, 108) = .44, $p > .77$) as was the main effect of group (F (1, 54) = .14, $p > .87$).

Second, when comparing the artificial superordinate-superordinate group with the meaningful superordinate-superordinate group at block 5, results showed a main effect of item type (F (2, 108) = 7.50, $p < .001$). The interaction between item type and group was statistically non-significant (F (2, 108) = .01, $p > .99$) as was the main effect of group (F (1, 54) = .02, $p = .95$).

Finally when comparing the artificial superordinate-superordinate group with the meaningful superordinate-superordinate group at block 9, results showed a main effect of item type F (2, 108) = 8.23, $p < .001$. The interaction between group and item type was statistically non-significant (F (2, 108) = .05, $p > .95$) as was the main effect of group (F (1, 54) = .67, $p > .73$).

The next sets of analyses compare artificial and meaningful label groups for prototype, fifty-fifty, and standard items, on blocks 1, 5, and 9, for basic level categories.

First, when comparing the artificial basic-basic phase 1 condition with the meaningful basic-basic phase 1 conditions, results showed a main effect of item type (F (2, 108) = 5.55, $p < .10$) and a significant main effect of group (F (1, 54) = 9.01, $p < .001$). The interaction between item type and group was statistically non-significant (F (2, 108) = .33, $p > .68$). The *artificial group* was faster than meaningful group when prototype (t (54) = 9.73, $p < .003$), and fifty-fifty items (t (54) = 4.87, $p < .05$) were classified at block 1, and when standard items were classified at block 1 (t (54) = 11.14, $p < .007$) and 2 (t (54) = 4.66, $p < .03$).

Second, when comparing the artificial basic-basic phase 2 condition with the meaningful basic-basic phase 2 condition at block 5 results showed a significant main effect of item type (F (2, 108) = 6.77, $p < .04$). The interaction between item type and group was statistically non-significant (F (2, 108) = 1.09, $p > .25$) as was the main effect of group (F (1, 54) = .03, $p > .90$).

Finally, when comparing the artificial basic-basic phase 2 condition with the meaningful basic-basic phase 2 condition at block 9, results showed a significant main effect of item type (F (2, 108) = 5.77, $p < .01$). The interaction between item type and group was statistically non-significant (F (2, 108) = 1.77, $p > .12$) as was the main effect of group (F (1, 54) = .01, $p > .99$).

In sum, no statistical differences were found between groups for items when categorizing at superordinate levels. However, faster performance was found for the artificial label group when processing standard, prototype and fifty-fifty items for basic level categories. These findings would suggest that mapping between item information and the category is poorer for the meaningful group during initial stages of learning.

5.5.2 Item comparisons for individual groups

Analyses for Experiment 4 revealed accuracy differences between the three items, but minimal differences in reaction times. This failure to find reaction time differences may have resulted from performance asymptote by block 9. Introducing prototype and fifty-fifty items into block 1 and 5 has the advantage of examining this possibility. An additional advantage follows from closer examination of transfer patterns for prototype and fifty-fifty items.

5.5.2.1 Artificial label group

Means and mean square errors for accuracy are depicted in figures 21, 25, and 29 and for response times in figures 23, 27, and 31. Because there were no logical differences between superordinate-superordinate and superordinate-basic groups or between the basic-superordinate and the basic-basic groups at block 1, these groups were combined for present analyses. Accuracy data are explored first, followed by response times.

5.5.2.2 Accuracy for block 1

Prototype items were classified better than standard items in the superordinate condition ($t\ (27) = 5.22$, $p < .001$) and the basic condition ($t\ (27) = 3.54$, $p < .001$). No statistical differences were found between prototype and fifty-fifty item in the superordinate condition (t (27) = .56, $p > .77$). Finally, fifty-fifty items were classified better than standard items in the superordinate condition (t (27) = 7.88, $p < .001$).These findings are consistent with the previous experiment and suggest that participants categorize based on the number of statistical values that correctly predict the category.

5.5.2.3 Accuracy for block 5

. Prototype items were classified better than standard items in both the superordinate-superordinate phase 2 (t (27) = 5.00, $p < .001$) and the basic-superordinate phase 2 condition (t (27) = 10.89, $p < .004$). Second, prototype items were classified better than fifty-fifty item in the basic-basic phase 2 condition (t (27) = 7.55, $p < .001$) and the superordinate-basic levels phase 2 condition (t (27) = 3.79, $p < .001$). No differences were found between prototype and fifty-fifty items in either the superordinate-superordinate phase 2 condition (t (27) = .01, $p > .99$) or the basic-superordinate levels phase 2 condition, (t (27) = .44, $p > .65$). Fifty-fifty items were classified better than standard item in the superordinate-superordinate phase 2 (t (27) = 5.01, $p < .001$) and the basic-superordinate levels phase 2 conditions (t (27) = 5.91, $p < .001$).

The failure to find differences between prototype and fifty-fifty items immediately following taxonomic transfer from the basic to superordinate level is important. As noted earlier, given an effect of taxonomy, these items would most likely differ at the point of taxonomic transfer. Particularly in the basic-super, because participants have just been taught categories that corresponded to prototype items better than to fifty-fifty items.

5.5.2.4 Accuracy for block 9

At block 9 prototype items were classified better than standard items in both the superordinate-superordinate phase 2 (t (27) = 4.63, p < .001) and basic-superordinate levels phase 2 conditions (t (27) = 7.00, p < .001). Prototype items were also classified better than standard items at both the basic-basic phase 2 (t (27) = 8.01, p < .001) and the superordinate-basic phase 2 conditions, (t (27) = 7.00, p < .001). Prototype and fifty-fifty items did not differ in either the superordinate-superordinate phase 2 (t (27) = .04, p > .97) or the basic-superordinate phase 2 conditions, t (27) = 1.66, p > .10. Fifty-fifty items were classified better than standard at both superordinate-superordinate (t (27) = 4.01, p < .005) phase 2 level, and basic-superordinate levels phase 2 level (t (27) = 4.90, p < .001).

5.5.2.5 Reaction times for block 1

Prototypes were classified faster than standard items in the basic condition (t (27) = 3.65, p < .001). However no differences were found between these items in the superordinate condition, (t (27) = 1.77, p > .17). No differences were found between prototype and fifty-fifty items in the superordinate condition (t (27) = .43, p > .55). Finally, no statistical differences were found between fifty-fifty items and standard items in the superordinate level (t (27) = 1.70, p > .06).

In sum, the only reaction time differences found occurred at the basic level with prototype items being classified faster than standard items.

5.5.2.6 Reaction times for blocks

No differences were found between prototype and standard items in either the superordinate-superordinate phase 2 (t (27) = 1.10, p > .34) and the basic superordinate levels phase 2 conditions (t (27) = .09, p > .90). Prototype items were classified faster than standard items in the basic-basic phase 2 condition (t (27) = 3.90, p < .003) but not in the superordinate-basic phase 2 condition (t (27) = 1.07, $p > .67$). No significant differences were found between prototype and fifty-fifty items in the superordinate-superordinate phase 2 (t (27) = .88, $p > .20$) and the basic-superordinate phase 2 condition (t (27) = 1.01, $p > .05$). Finally, fifty-fifty items were classified faster than standard items in the superordinate-superordinate phase 2 (t (27) = 3.98, $p < .001$) and the basic-superordinate phase 2 conditions (t (27) = 2.10, $p < .03$).

In sum, prototype items were classified faster than standard items at the basic-basic phase 2 condition. Moreover, fifty-fifty items were classified faster than standard items in the superordinate-superordinate phase 2, and the basic-superordinate phase 2 conditions.

5.5.2.7 Reaction times for block 9

Prototype items were classified faster than standard items in the superordinate-superordinate phase 2 (t (27) = 2.20, $p < .05$) and the basic-superordinate phase 2 conditions (t (27) = 2.56, $p < .05$). No significant differences were found between prototype and standard items in either the basic-basic phase 2 (t (27) = 1.34, $p > .11$) or the superordinate-basic phase 2 conditions (t (27) = 1.78, $p > .10$). No differences were found between prototype and fifty-fifty items in either the superordinate-superordinate phase 2 (t (27) = .04, $p > .95$) and basic-superordinate phase 2 conditions (t (27) = .45, $p > .67$). Fifty-fifty items were classified faster than standard items in both the superordinate-superordinate (t (27) = 2.30, $p < .07$) and basic-superordinate conditions, t (27) = 3.66, $p < .001$.

In sum, no response time differences were found between prototype and fifty-fifty items, however response times for these items were faster than for standard items.

5.5.3 Meaningful label group.

Means and mean square errors for accuracy are depicted in figures 22, 26, and 30, and for response time in figures 24, 28, and 32. Because there were no logical differences between superordinate-superordinate and superordinate-basic groups or between the basic-superordinate and the basic-basic groups these groups in block 1, these groups were combined for analysis. Accuracy data are explored first, followed by response times.

5.5.3.1 Accuracy for Block 1.

First, prototype items were classified better than standard items in both the superordinate (t (27) = 9.60, $p < .001$) and basic conditions (t (27) = 4.72, $p < .001$). No statistical differences were found between prototype and fifty-fifty items in the superordinate condition (t (27) = .66, $p > .30$). Fifty-fifty items were classified better than standard items in the superordinate condition (t (27) = 6.33, $p < .001$).

In sum, prototype and fifty-fifty items were classified better than standard items, however no differences were found between fifty-fifty items and prototype items. Thus, findings do not support the idea that participants viewed fifty-fifty items as odd otherwise performance for these items would have been poorer than for other items. Instead, findings suggest that participants classified based on the number of dimensional values that correctly predicted the category. Other explanations are explored in the discussion for this experiment.

5.5.3.2 Accuracy for Block 5

Prototype items were classified better than standard items in the superordinate-superordinate phase 2 (t (27) = 7.22, $p < .001$) and the basic-superordinate phase 2 conditions (t (27) = 9.77, $p < .001$). Prototypes were also classified better than standard items in both the basic-basic phase 2 (t (27) = 6.55, $p < .001$) and the superordinate-basic phase 2 conditions (t (27) = 5.24, $p < .001$). No differences were found between prototype and fifty-fifty items in either the superordinate-superordinate phase 2 (t (27) = .90, $p > .44$) or the basic-superordinate phase 2 conditions (t (27) = .88, $p > .45$). Fifty-fifty items were classified better than standard items in both the superordinate-superordinate phase 2 (t (27) = 4.22, $p < .001$) and the basic-superordinate phase 2 conditions (t (27) = 8.99, $p < .001$).

In sum, findings for block 5 replicate those of block 1, prototype and fifty-fifty items were classified better than standard items, however no differences were found between fifty-fifty items and prototype items. An important finding here is the failure to find differences between prototype and fifty-fifty items immediately following taxonomic transfer from basic to superordinate level. As noted earlier, taxonomic transfer is the point at which differences between these items was most likely to occur.

5.5.3.3 Accuracy for block 9

Prototype items were classified better than standard items in both the superordinate-superordinate phase 2 (t (27) = 11.00, $p < .001$) and the basic-superordinate phase 2 conditions (t (27) = 7.77, $p < .001$). Prototypes were classified better than standard items (t (27) = 9.90, $p < .001$) in both the basic-basic phase 2 (t (27) = 5.88, $p < .001$) and the superordinate-basic phase 2 conditions. No differences were found between prototype and fifty-fifty items when classified in either the superordinate-superordinate phase 2 (t (27) = .99, $p > .32$) or the basic-superordinate phase 2 conditions (t (27) = .55, $p > .56$). Fifty-fifty items were classified better than standard items in both the superordinate-superordinate phase 2 (t (27) = 7.23, $p < .001$) and the basic-superordinate phase 2 levels (t (27) = 6.01, $p < .001$).

5.5.3.4 Reaction times for block 1

First, prototypes items were classified faster than standard items at both superordinate (t (27) = 3.22, $p < .004$) and basic conditions (t (27) = 2.55, $p < .05$). No differences were found between prototype items and fifty-fifty items in the superordinate condition (t (27) = .03, $p > .98$). Finally, fifty-fifty items were classified faster than standard items in the superordinate condition (t (27) = 2.80, $p < .01$).

In sum, the finding of faster reaction time for prototype and fifty-fifty items over standard items coupled with the failure to find differences between prototype and fifty-fifty items suggests that participants classified based on the number of dimensional values belonging to categories.

5.5.3.5 Reaction times for block 5

No reaction time differences were found between prototype and standard items in either the superordinate-superordinate phase 2 (t (27) = 1.85, $p > .09$) and basic-superordinate phase 2 conditions (t (27) = 1.11, $p > .40$). Prototypes items were classified faster than standard items in the superordinate-basic phase 2 (t (27) = 2.40, $p < .05$) and the basic-basic phase 2 conditions (t (27) = 3.78, $p < .001$). No differences were found between prototype and fifty-fifty items when classified in either the superordinate-superordinate phase 2 (t (27) = 1.01, $p > .05$), and basic-superordinate phase 2 conditions (t (27) = .24, $p > .80$). Fifty-fifty items were classified faster than standard items (t (27) = 5.01, $p < .001$) in the superordinate-superordinate phase 2 condition, but not in the basic-superordinate phase 2 condition (t (27) = .40, $p > .71$).

5.5.3.6 Reaction times for block 9

Prototype items were classified faster than standard items in the superordinate-superordinate phase 2 condition (t (27) = 3.00, $p < .01$) but not in the basic-superordinate phase 2 condition (t (27) = .22, $p > .65$). Prototype items were classifid faster than standard items in both the basic-basic phase 2 (t (27) = 2.01, $p < .05$) and the superordinate-basic phase 2 conditions (t (27) = 2.60, $p < .01$). No differences were found between prototype and fifty-fifty items when classified in either the superordinate-superordinate phase 2 (t (27) = .78, $p > .54$) or the basic-superordinate phase 2 conditions (t (27) = .65, $p > .53$). Fifty-fifty items were classified faster than standard items in the superordinate-superordinate phase 2 condition (t (27) = 4.44, $p < .001$) but not in the basic-superordinate phase 2 condition (t (27) = 1.07, $p > .35$).

In comparisons to standard items, participants were generally faster processing prototype and fifty-fifty items. The one exception was the failure to find a difference between these items for the basic-superordinate phase 2 condition. It would seem that standard items benefited more from basic level priming than other items. Indeed, reaction times for these items were faster when transferring from basic to superordinate levels than when transferring from superordinate to superordinate levels. This finding may be limited to data set, as similar findings were not found in Experiment 4.

5.5.4 Differences between taxonomic levels

The next sets of analyses examine mean response time differences between superordinate and basic level categories for artificial and meaningful groups. These analyses are important for differentiating the influence of prior expectations, dimensions, and item structure, on categorization. For example, slower responses on part of the meaningful group for prototype items would suggest that abstract dimensions negatively impacted performance. This is because the primary factor affecting performance for prototype items is abstract dimensions (structure for prototype items was held constant between taxonomic levels). However, slower responses to standard items on part of the meaningful group would suggest that both abstract dimensions and incongruent dimensional values negatively affected meaningful group performance (these items have both abstract dimensions and incongruent dimensional-values). Finally, slower between taxonomic level responses on part of the meaningful group when comparing standard and prototype items on would suggest that incongruent dimensional values are the primary factor affecting performance. The first set of analyses compares mean differences between basic-basic and superordinate-superordinate groups. The second set of analyses examines group mean differences between basic-superordinate and superordinate-basic phase 2 conditions.

Results were inconclusive when comparing differences between taxonomic levels for meaningful and artificial groups in Experiment 4. Findings showed that the artificial group was faster processing standard items, however no differences were found between groups for prototype items. Thus, it was difficult to determine with any certainty to what extent abstract or incongruent dimensional values contributed to findings. One reason for failing to find an effect of abstract dimensions for prototype items may follow from items' characteristics being less surprising to the meaningful group by block 9. Given this possibility one would expect greater differences between groups when processing prototype items at block 1. For similar reasons one might also expect that given an effect of incongruent dimensional values, differences between taxonomic levels when comparing prototype and standard items would also be greater for the meaningful label at block 1.

First, when comparing differences between basic-basic and superordinate-superordinate groups at block 1, mean differences were smaller for the artificial label when categorizing standard (t (54) = 2.44, $p < .01$) and prototype items (t (54) = 3.40, $p < .004$). No statistical differences were found between groups when categorizing items on blocks 5 ($p > .36$) and 9 ($p > .82$). These findings support the idea that abstract dimensions interfere with the meaningful groups mapping of item information. Second, comparing differences between basic-superordinate and superordinate-basic phase 2 condition, differences were smaller for the artificial group when categorizing standard items at block 5 (t (54) = 2.67, $p < .05$). All other comparisons were statistically non-significant ($p > .70$).

Finally, the next comparison explores the idea that incongruent dimensional values interfere with mapping of item information. First, focusing on block 1, although observable mean differences between standard and prototype items were greater for the meaningful group (M = 1579) than the artificial group (M = 663) findings were statistically non-significant (t (54) = 1.00, $p > .77$). Comparisons were also statistically non-significant, at blocks 5 (t (54) = .88, $p > .95$) and 9 (t (54) = .82, $p > .91$). Thus, at least statistically an effect of dimensional values goes unsupported.