12-4-2006

What Meaning Means for Same and Different: A Comparative Study in Analogical Reasoning

Timothy M. Flemming

Follow this and additional works at: http://scholarworks.gsu.edu/psych_theses

WHAT MEANING MEANS FOR SAME AND DIFFERENT:

A COMPARATIVE STUDY IN ANALOGICAL REASONING

by

TIMOTHY M. FLEMMING

Under the Direction of David A. Washburn

ABSTRACT

The acquisition of relational concepts plays an integral role and is assumed to be a prerequisite for analogical reasoning. Language and token-trained apes (e.g. Premack, 1976; Thompson, Oden, and Boysen, 1997) are the only nonhuman animals to succeed in solving and completing analogies, thus implicating language as the mechanism enabling the phenomenon. In the present study, I examine the role of meaning in the analogical reasoning abilities of three different primate species. Humans, chimpanzees, and rhesus monkeys completed relational match-to-sample (RMTS) tasks with either meaningful or nonmeaningful stimuli. For human participants, meaningfulness facilitated the acquisition of analogical rules. Individual differences were evident amongst the chimpanzees suggesting that meaning can either enable or hinder their ability to complete analogies. Rhesus monkeys did not succeed in either condition, suggesting that their ability to reason analogically, if present at all, may be dependent upon a dimension other than the representational value of stimuli.

INDEX WORDS: Analogical reasoning, Analogies, Meaning, Relational concepts, Same/different, Nonhuman primates

WHAT MEANING MEANS FOR SAME AND DIFFERENT:

A COMPARATIVE STUDY IN ANALOGICAL REASONING

by

TIMOTHY M. FLEMMING

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2006

WHAT MEANING MEANS FOR SAME AND DIFFERENT:

A COMPARATIVE STUDY IN ANALOGICAL REASONING

by

TIMOTHY M. FLEMMING

Major Professor: David A. Washburn
Committee: Michael J. Beran
Eric J. Vanman
Heather M. Kleider
Roger K. R. Thompson

Electronic Version Approved:

Acknowledgments

First and foremost, I thank my adviser and thesis chair, Dr. David A. Washburn, for his unending support on this and other projects. His advice has been integral to the success of this and many other projects. I also thank my undergraduate adviser and committee member, Dr. Roger K. R. Thompson for kindling my interest in this area of research. I thank my additional committee members, Dr. Michael Beran, Dr. Eric Vanman, and Dr. Heather Kleider for their insightful comments both on written drafts and throughout the development of this project. In addition, I thank Dr. Beran for his programming assistance on software used in this study. I also thank Ted Evans, Sarah Hunsberger, and John Kelley for their assistance in data collection. Lastly, special thanks are given to Dr. Washburn, Dr. Thompson, Dr. Beran and Emily Harris for providing constant reminders that pursuit of a failing project can be more rewarding than finding what we expected.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

Conceptual thinking affords us the opportunity to make sense of and to organize our environment in such a way that it is more meaningful and far more manageable. We are predisposed to organize our field of view into meaningful groups and search for similarities among them, thus enabling them to be useful in our everyday lives. Of interest to comparative psychologists is the degree to which our nonhuman primate relatives posses similar capabilities and find such capabilities helpful in their everyday lives. The range of conceptual abilities has been extensively investigated for several species of the animal kingdom, including birds (Pepperberg, 1987; Cook, Wright, & Kendrick, 1990; Wasserman, Young, & Fagot, 2001; Wright, Rivera, & Katz, 2003), dolphins (Herman, Hovancik, Gory, & Bradshaw, 1989), monkeys (Bovet & Vauclair, 2001; Flemming, Beran, & Washburn, 2005; Katz, Wright, & Bachevalier, 2002; Shields, Smith, & Washburn, 1997), and apes (Premack, 1976; Thompson & Oden, 1996). But, how far are nonhuman animals able to abstract these conceptual abilities in order to apply them to novel situations? One reason that such a question is intriguing is that the lower levels of conceptual abilities observed in several nonhuman species afford us the unique opportunity to examine the continuity of these cognitive processes across humans and other animals (Herrnstein, 1990).

A concept is a set of characteristics shared by all those and only those instances in a particular set (Engelmann, 1969). Concepts provide a means to solve new problems in novel situations. Without the ability to think in terms of concepts, we might be overwhelmed by the complexity of our environments (Bruner, Goodnow, & Austin, 1956). Rather than searching for differences between objects that we encounter, concepts

serve as instruments for simplifying our surroundings by determining how things are *alike* (Herrnstein, 1990; Pearce, 1994; Zentall, Galizio, & Critchfield, 2002). Numerous species including pigeons, monkeys, and humans show an affinity toward identity; these species show a great deal of proficiency at matching *same*, but not *different* displays (Fagot, Wasserman, & Young, 2001; Wasserman, Frank, & Young, 2002 Young and Wasserman, 2002).

Conceptual abilities, like many other cognitive processes, lie on a continuum from basic and perceptual to more abstract and even relational. Thomas (1980) proposed a hierarchy of learning-intelligence that includes several levels of conceptual ability. Levels 1-5 of Thomas's learning-intelligence hierarchy outline basic habituation, signal conditioning (Pavlovian conditioning), stimulus-response (operant) learning, chaining (chaining stimulus-response learning units), and concurrent discrimination learning, respectively. Level 6 concerns class concepts, the most basic level of conceptual representation (Steirn & Thomas, 1990). At this level, objects can be processed together because of their perceptual similarities with one another. In addition, more abstract mental representations for classes of objects can be created and applied to arrays of objects in order to group them together. For example, writing instruments such as pencils and computer keyboards may share no physical similarities with one another, but can nonetheless be grouped together because they afford the same functionality to the user.

These concepts can be applied further to relational concepts, also referred to as abstract relations (Herrnstein, 1990; Thompson & Oden, 1996) as described in levels 7 and 8 of Thomas's (1980) learning-intelligence hierarchy. At these levels, classification deals not with the exemplars themselves, but with the relations between and among

concepts (i.e., sameness and difference). Rather than attending to discrete physical

qualities, we can examine the relatedness of one object to another and decide how they

are related (Premack, 1976; Thompson, 1995). Perhaps the highest order of conceptually

mediated behavior is the ability to judge relations-between-relations that forms the

necessary foundation for analogical reasoning, which many regard as the hallmark of

human reasoning and intelligence (Gentner & Markman, 1997; Holyoak & Thagard,

1997; Premack, 1983; Sternberg, 1977).

Studying Concepts in Nonhuman Animals

Certainly for nonhuman animals, conveying functional definitions of concepts

cannot be accomplished through verbal communication. Rather, we must rely on

nonverbal conceptually mediated behavior. Just as we teach concepts to children who

understand spoken language, we require an analog for studying such conceptual abilities

in nonhuman animals. One way to demonstrate the understanding of a concept is through

nonlingual overt behaviors. Conceptually mediated behavior is both functionally adaptive

and cognitively efficient in ways that it permits animals to judge and to adjust their

behavior to novel objects and events by virtue of membership in an *already* familiar class

(Cook et al., 1990; Herrnstein, 1990; Premack, 1983; Wasserman, et al., 2002. By

making these explicit similarity judgments both between and within common classes of

objects, animals can convey understanding of conceptual information without the use of

spoken language. The animal's behavior can thus be said to reflect their conceptual

understanding (Pearce, 1994). This said, it is difficult for us to know precisely how

animals would define their concepts, but certainly not impossible.

Roberts and Mazmanian (1988) investigated concept acquisition at varying levels of abstraction in three different species: humans, pigeons, and squirrel monkeys. Utilizing concepts at three levels of abstraction similar to those described above (Thomas, 1980), Roberts and Mazmanian employed a two-choice discrimination task that required animals to differentiate one slide from another. Slides were presented two at a time to the animals. Animals pressed a key corresponding to one of the slide choices on any given trial. If a slide was deemed to be "in-category" selection of it was rewarded with either a light (humans), a banana pellet (monkeys), or the brief presentation of a grain hopper (pigeons). If the chosen slide was not "in-category" no reward was dispensed. Slides were photographs of animals from an assortment of books and magazines. Slides with animals contained a wide variety of species including insects, fish, birds, amphibians, reptiles, and mammals. A disproportionately larger number of pictures came from the bird class, as the basic level discrimination was that of the common kingfisher (*Alcedo atthis*) from other birds. All pictures varied in the viewpoint from which the photograph was taken, as well as coloring, number of animals in the picture, and proportion of animal to background area. The pictures not containing animals included a variety of indoor and outdoor scenes with trees, flowers, mountains, foods, clothing, airplanes, and houses. After 30 sessions of training with one set of slides, eight days of transfer testing with probe stimuli were conducted.

Humans, not surprisingly, were able to acquire concepts at the basic, low-abstraction, as well as the high-abstraction levels. Humans correctly chose the "in-category" slides with around 90% accuracy for all three levels. Monkeys and pigeons, however, were less successful at certain levels of abstraction. Monkeys were significantly

better at making the discriminations at low (kingfisher vs. other bird) and high (animal vs. nonanimal) levels. Pigeons only successfully acquired the most basic concept: they discriminated only kingfishers from nonkingfishers. All three species appeared to have formed the kingfisher concept. When the problem was made more abstract by requiring subjects to identify birds in general, or animals in general, the category may have become too broad or abstract for the subjects to learn a simple rule for identification for identifying individual exemplars (Roberts & Mazmanian, 1988). These findings support the theory that animals learn concepts by responding to a small set of features in pictures that look similar (Premack, 1983; Zentall et al., 2002).

When studying concept formation and acquisition in our own species, we can simply ask participants about the rules they used in order to perform successfully, although research shows that these self-reports may not be veridical (Gentner & Markman 1997). However, one can be successful at some categorization tasks without ever acquiring a concept that is defined in the same way as it may be by the majority of humans. In an experiment by Savage-Rumbaugh, Rumbaugh, Smith, and Lawson (1980) chimpanzees were required to sort a mixed pile of tools and food into two separate piles based on these categories. Apes were trained with a small set of objects and were successfully able to sort new, but familiar, objects during test trials. It is not easy to argue that this problem was solved on the basis of physical similarity of the test items to the training items, because it is difficult to identify a set of physical features that an object must possess in order to be classified as a tool or as food. Instead, these objects may have been categorized successfully because the subject possessed the concepts "food" and "tool." Conversely, these objects may not have been categorized utilizing the same

meaning to the concepts that we might have. For instance, the chimpanzees may use just one concept of "food/edible" meaning to them "something I would or have in the past eaten" and "not food/inedible" for the tools. This type of sorting does not prove that the animals have indeed acquired a "tool" concept. Although we could be using the same strategy with similar names for the concepts that differentiate the classes of objects, we tend also to devise a unique and more specific concept for sets of objects that describe how they are related to each other, rather than simply how they are unlike those objects that fall under another concept (Savage-Rumbaugh et al., 1980).

Like chimpanzees sorting slides into tool and. food groups, capuchin monkeys have proven proficient at sorting pictures into person/nonperson groups. In the case of capuchin monkeys sorting photos based on the "person" concept (D'Amato & Van Sant, 1988), categorization may mimic how humans view the person concept. D'Amato and Van Sant trained capuchin monkeys to categorize photographs based on the "human" concept. Monkeys discriminated between slides containing humans from those not containing human figures. Although monkeys were successful on the task, D'Amato and Van Sant found several interesting errors by analyzing individual test trials. Any nonperson slide that contained a red patch (i.e., other animals or fruits) was more likely to be classified as belonging to the person category. It is therefore possible that a "red patch" was a feature that the monkeys determined sufficient for responding and that it acquired considerable associative strength, presumably because it is common to many faces and resulted in a high level of responding whenever it was presented. The primary way we know that humans use a true "person" concept rather than a "red patch" concept is by asking them.

Bovet and Vauclair (2001) similarly investigated abstract concept formation in baboons. Food and nonfood objects were presented to the monkeys with a modified Wisconsin General Test Apparatus (WGTA). Baboons were trained to pull one of two ropes corresponding to same or different (S/D) in response to the relation between the two objects presented. If both objects presented were from the same category (food/nonfood), pulling the rope corresponding to *same* was the correct response. When one food and one nonfood were presented simultaneously, pulling the *different* rope was the correct response. Although the authors suggested that success on the part of the monkeys to complete the task was adequate demonstration of judgment of conceptual identity, the results do not imply any sense of relational matching. Making one response for two food objects versus one food and one nonfood does not imply that the monkeys have any S/D concept. Rather than implying conceptual labels, the response ropes could simply symbolize presence or absence of food.

Relational Discriminations

Just as basic concepts act as simplifiers of our world by allowing us to group objects into categories, abstract concepts provide a means to extend this function beyond physical or functional similarities (Pearce, 1994). Subjects can choose between discriminative stimuli based not only on absolute physical properties, but also on *relations between* those stimuli that are presented.

Kinnaman (1902) was one of the first to recognize the distinction between types of discrimination, suggesting that alternative strategies exist to individuals solving these kinds of tasks. They may rely more on the relative than the absolute properties of the stimuli. For example, when given a choice between a 4 cm tall box and a 1 cm tall box, a

subject may first begin choosing under the assumption that the rule is to choose the 4 cm tall box (based on absolute size). However, when more examples are given, it becomes apparent that absolute size is not the rule under which the paradigm operates. Rather, one must compare the two boxes presented and choose the larger of the pair (relative size), whether it is 1 cm or 4 cm tall. That is, the object that is larger relative to the other object.

Kirkpatrick-Steger and Wasserman (2000) investigated pigeon's abilities to discriminate among stimuli using relative information of shape and location. Pigeons were trained to peck a pairs of shapes that were arranged in one of several different configurations in relation to each other. Some items were, for example, "to the right of" or "on top of" the second item of the pair. Rather than attending to the specific shapes included in the sample, pigeons were required on some trials to attend specifically to the relation of one object relative to the other. Whereas some pigeons were unable to overcome the salience of shape, others succeeded in learning relative configural rules.

Analogical Reasoning

Analogical reasoning, Halford and Graeme (1992) argued, is the mechanism that allows for all conceptual thinking, including logical inference. Knowledge about analogies forces explicit expression of conceptual knowledge, unlike simple discriminations that may rely on more implicit types of conceptual knowledge (Premack, 1986). In an analogy, a relationship must be established between the first two elements in the series. Then, and only then, can one continue to the second set of elements and seek the same relation between them. By discriminating between two abstract relations, one is able to acquire the knowledge needed to complete and construct analogies, much like the chimpanzee, Sarah (Premack, 1983). Sarah was given a variety of analogical reasoning

problems using arrays of meaningful plastic chips of different colors and shapes. Two

tangible plastic objects that varied on one dimension (i.e., color, shape, or size) were

placed to the left of a center chip which signified *same*. To the right of the *same* symbol

was placed only one object. The task thus required the chimpanzee to perceive the

relationship between the shapes on the left and recreate its analog to the right of the

center chip. Not only was Sarah able to complete the task with flat geometric shapes, but

she was also successful when the items presented were everyday three-dimensional

objects (Gillan, Premack, & Woodruff, 1981).

Along the primate lineage, species more closely related to humans have many

common cognitive abilities, including the capacity for judging relations-between-

relations through the further application of conceptual knowledge. Many ape species,

including humans, are therefore considered "analogical" because of their adept ability to

represent the world propositionally. The fundamental distinction between monkeys and

apes is in their explicit conceptual capacities. Monkeys accept identity based upon

identical *features*, whereas humans (and other ape species) accept it on the basis of

identical *subject matter*. Monkeys, therefore, can be thought of as "paleo-logicians" in

the sense that they form common class concepts of identity on the basis of common

physical features. This is not because monkeys do not represent their world, but rather

because they appear not to represent it *propositionally* (Thompson & Oden, 2000). Rather

than recalling representations of the *meaning* expressed in what they see, monkeys tend

to represent their environments exactly as they are expressed.

Some abstract relations can be visually perceived at some level and are

subsequently easier to process. Abstract perceptual relations can be visualized, but

require some level of abstract conceptualization, possibly at a lesser degree, compared to other abstract relations. Fortes, Merchant, and Georgopoulos (2004) investigated spatial judgments in animals by evaluating the ways in which the rhesus monkey acquired the concepts of *high* and *low*. Differentiating between *high* and *low* is a rather easy task for human adults. In a delayed matching-to-sample task, a rhesus monkey determined whether lines on a computer screen should be categorized as *high* or *low*. Beginning with the extremes (top and bottom) of the computer monitor, bars were displayed at varying heights. After each bar was presented for one second, choice stimuli appeared (similar bars at the extreme top and bottom of the screen) allowing the monkey to respond *high* or *low*. Throughout a testing session, bars varied in height, with many hovering around the midline of the screen. As for humans, these bars proved more difficult to classify by the monkey. The monkey was able to generate an abstract notion of a midline in a fashion similar to adult humans. This evidence supports the notion that nonhuman primates can generate an abstract *perceptual* concept.

In addition to examining the absolute acquisition of such concepts, Fortes, Merchant, and Georgopoulos (2004) sought to determine how the concepts are represented in nonhuman animals. In humans, similar concepts may be represented either semantically, or by an analog comparison model. One way that this can be examined is by the existence of the congruity effect. The congruity effect occurs when there is a decrease in the response time when the objects compared are closer to the category pole. In this instance, when bars are closer to the midline, response time may increase. Upon determination of the existence of the congruity effect in the rhesus monkey, an analog comparison is more likely to be favored by nonhuman animals (Fortes et al., 2004).

Same and Different

The judgment of *same* versus *different* (S/D) is such an exceedingly straightforward one for humans that it often becomes difficult to make comparisons and describe events without these words. For this reason, Premack and Premack (2003) called *same* and *different* "early concepts". When making comparisons between objects or events in the world, the same-different distinction is many times our first approach, conveying the most useful information (Premack, 1976). Responding differentially to groups of same or different items seems an almost trivial task. For humans, however, the S/D rule of discrimination is one that we have come to rely on, and tend to use readily. Making relational judgments, although simple and obvious to adult humans, may not be a simple and obvious default rule for discrimination groups of objects for other species (Premack & Premack, 2003).

Whereas it is often difficult to describe relations without using the words *same* and *different*, the abstract concepts have no linguistic prerequisites (Premack, 1976). The words can be applied to objects that themselves do not have names. In addition, when using relation-level problem solving, the name of the objects is irrelevant. Rather, the relation *between* the objects is the only relevant information. In relational matching tasks, an individual must abandon ordinary matching entirely and move to a different level of problem solving--the relational level (Premack & Premack, 2003). To complete a relational match-to-sample successfully, participants must direct questioning to "what is the relationship between the members of each pair?" Labeling of the pairs is the necessary component to any relational match-to-sample task. So, participants must match self-given labels of pairs rather than the pairs themselves. Dependence on labels brings

to the forefront the question of whether language may be necessary to make judgments of

relations between relations (Thompson, Oden, & Boysen, 1997). Can we sufficiently

label objects without language? If the answer to this question is "no," then we should not

be surprised by Premack's suggestion that participants, human or nonhuman, cannot

complete a relational matching task without extensive language training. However, if

effective labeling for relational information can be accomplished without language *per

se*, then we should expect language-naïve nonhuman species also to succeed on tasks

requiring analogical thinking.

Throughout recent years, several studies have indicated that Premack might have

been right: language training is necessary when relational problem-solving skills are

required (Shyan & Wright, 1993; Thompson & Oden, 1996 Wright & Santiago, 1984).

Related research shows, however, that nonhuman animals such as pigeons and monkeys

possess implicit knowledge of the concepts of same and different, no matter how limited

and different from our own knowledge of the concepts (Fagot et al., 2001; Wasserman et

al., 2002; Wright et al., 2003).

Two types of studies comprise a majority of the S/D concept literature:

acquisition of the concepts themselves and their application to the judgment of abstract

relations. To determine the acquisition of the concepts themselves, subjects typically

need only to respond differentially to groupings of exemplars that are either all identical

or all different. When applying the judgments to abstract relations, subjects often must

use the concepts as the basis for their future behavior and match one set of exemplars to

another set on the basis that the relation is the same for the sets.

The matching concept itself is abstract (Wright, 1997). It is abstract because it transcends the stimuli used to train it and is distinct from "natural" concepts which are categories unified by some specific stimulus attribute or attributes. In the traditional match-to-sample paradigm, pigeons (Wright, 1997) and rhesus monkeys (Washburn, Rumbaugh, & Richardson, 1992; Washburn & Astur, 1998) were capable of responding on the basis of the simple if/then statement: "*if* sample equals A, *then* choose A not B." However, when tested on novel stimuli in transfer tests, pigeons often failed the task. They had learned the configural patterns necessary for performance on training tests, but had not truly learned the concept necessary to succeed on transfer tests. That is, they do not typically learn the matching concept. The reason that pigeons may learn configural patterns instead of abstract concepts may reflect their learning predisposition (Wright, 1997). Concept learning requires that subjects learn to relate one stimulus to another – to relate each comparison stimulus to the sample stimulus.

Cook, Katz, and Cavoto (1997) examined the acquisition of the S/D concepts through the use of a two-choice discrimination task by pigeons. Birds were presented displays from four distinct domains classified as either the same or different: texture, feature, geometric, and object. The use of four different display types was integral in determining whether the birds truly acquired the concepts of same and different as we understand them so that they can be universally applied across different domains. *Same* displays consisted of the repetition of a single element throughout a 24 x 6 array. In the *different* displays, an 8 x 7 region of contrasting elements (differing in either shape or color) was randomly located within the larger array. Texture displays appeared uniform, with the exception of a nonidentical patch of simulated texture for different displays.

Geometric patters were arrays approximately six of the same geometric object, or arrays of five of the same and one odd object. Object displays were patterned in the same fashion, but were made of actual clipart pictures rather than simple geometric objects. Pigeons were tested on all four types of displays presented on a computer monitor. After 50 sessions of testing concurrently across all display types, pigeons had readily learned a same-different discrimination. The rate and general pattern of responding was approximately the same for all types of displays, lending support that the concepts are applied identically across different domains. A single generalized rule was most likely used to discriminate all display types. These results provided some of the first strong evidences that pigeons, like many primates, can learn and abstract, visually mediated S/D concept (Cook et al., 1997).

Shields, Smith, and Washburn (1997) presented monkeys with a S/D task in which the discrimination required was between boxes containing various amounts of pixilation. The main goal of this study was to examine the uncertainty response as it applies to judgments of same versus different. Pairs of pixel boxes were presented on a computer screen with a star (a previously used symbol for escape) and a "D." If the pixel boxes matched in their amount of pixilation (i.e. their "sparseness" or "denseness"), the correct choice was to move the cursor in contact with the box pair. If the boxes were illuminated with different amounts of pixels, choosing "D" was correct. As the density ratio between the two boxes approached 1.0, the same-different distinction became increasingly different, leading to a choice of the star to escape the trial. This evidence supports the notion that monkeys may understand the S/D distinction along a continuum.

After several training sessions, monkeys discriminated between *same* and *different* relations with a great deal of accuracy.

Identity and Nonidentity

Researchers have demonstrated that nonhuman animals can learn to discriminate large arrays in which all of the items are physically identical or nonidentical to each other (Wasserman et al., 2001). That is, they can detect perceptual similarities and differences. Evidence is insufficient, however, to suggest that their concepts of *sameness* and *difference* operate in a way that is congruous to our own. One caveat to the above mentioned studies is that the animals are unable to make such distinctions when the arrays contain fewer than 6-8 items, suggesting that subjects are actually relying on *variability* or *entropy* in order to succeed at the task. Rather than recalling a representation of *difference*, animals may be perceiving the amount of perceptual variance or regularity to be greater in the *different* arrays than in the *same* arrays (Fagot et al., 2001). When the number of items displayed decreases, we observe depreciation in the animals' ability to discriminate.

Entropy: Same vs. Different?

In a series of experiments with pigeons (Young & Wasserman, 1997) the effects of number of items in a given display as it relates to the rate of concept acquisition was investigated. After several animal species were shown to fail on tasks involving the categorization of just two visual items as the same or different (Santiago and Wright, 1984), Wasserman and colleagues devised a task that would make these types of tasks easier. Perhaps, they reasoned, there is simply not enough information in a two-item display to convey the concepts of *sameness* and *difference*. For this reason, Wasserman's

studies typically introduce animals to arrays of 16 or more items for discrimination (see Figure 1).

The notion of entropy describes the phenomena that might occur when success in these types of paradigms is contingent upon the number of items in a display. Entropy is described as the amount of change or variability that is perceived within a grouping of items. This idea was put forth by Shannon and Weaver (1949). To compute entropy one can use the following equation, where *H(A)* is the entropy of a categorical variable *A*, *a* is a category of *A*, and p_a is the proportion of observed values within that category:

$$H(A) = - \Sigma \, p_a \log_2 p_a$$

Thus, in a 16-item array of all identical items, entropy is independent of the number of items: entropy = 0 (for any set of identical items). Changing the number of exemplars in the sample display should not change the amount of pictured variability. For *different* displays, however, reducing the number of items in the sample would reduce the amount of pictured variability. With only two pictured items, the variability of *different* sample displays (entropy = 1) is numerically closer to 16-item *same* samples (entropy = 0) than it is to 16-item *different* samples (entropy = 4).

Pigeons proved successful in transferring their knowledge to novel stimuli, lending support that S/D concepts were truly formed. In addition, both species were strongly controlled by the entropy of sample displays. Both humans and baboons used the entropy of the sample to guide their choice behavior.

However, the key distinction between the species is where they set the entropy cutoff for these displays. It appears that humans set that cutoff near 0, whereas baboons set that cutoff near the midpoint range of the entropy scores, here approximately 2.
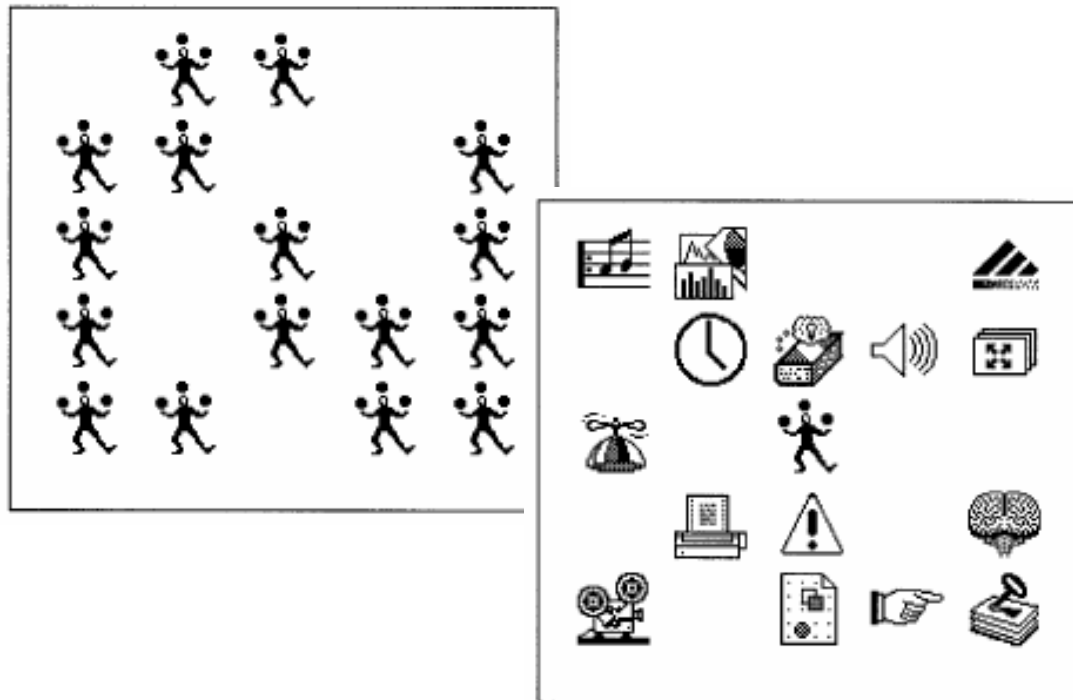
Figure 1.Entropy-Infused Displays Presented to Pigeons in Wasserman, Young, and Peissig, (2002).

Following the finding that baboons could also successfully differentiate between S/D 16-icon displays (Wasserman et al., 2001 in a two-choice task, discrimination of second order relations was examined. Fagot, Wasserman, and Young (2001) investigated whether baboons could discriminate *same* from *different* using their abilities to judge relations-between-relations in a delayed relational match-to-sample task (RMTS). For comparison purposes, adult humans were also required to complete the same task. 16-icon arrays were again used as stimuli, but were presented in the MTS format. One array appeared on the computer screen as the sample array, followed by a small delay and the presentation of two choice arrays, only one of the choice arrays being of the same relational type as the sample. Two baboons successfully learned the RMTS. That is, they accurately picked the choice display that involved the same relation among the 16 icons (*same* or *different*) as in the sample display. The acquisition of RMTS for baboons was comparatively slow. In the final four sessions after approximately 6,000 trials, baboons averaged 84% correct. In addition, success on *same* trials was acquired at a more rapid rate. In comparison, humans were approximately 100% accurate on both stimulus types within the first 100 trials. After continued training, it was determined that given enough trials (more than 10,000), one baboon successfully discriminated (81% correct) S/D displays of only 3 items.

In a second experiment, Wasserman, Young, and Fagot (2001) compared the effects of number of icons between species. As in previous studies, performance on the delayed MTS task dropped sharply if the number of icons in the display was reduced below 8 items (after which performance did not exceed 63%). In contrast, display set size

had no effect on adult human participants. Performance did not diminish as set size decreased for humans as it did for baboons.

Rather than training and requiring subjects to respond to abstract conceptual stimuli in order to obtain reward, implicit discriminations enable us to examine how animals may be predisposed to process stimuli. Wasserman, Frank, and Young (2002) asked whether pigeons might exhibit relational stimulus control that did not explicitly require S/D discrimination. As in previous experiments (Fagot et al., 2001; Wasserman et al., 2001), pigeons viewed arrays of icons that were either all the same as each other or all different from each other. In addition, arrays were composed of icons from one of two sets, creating four distinct array types: *different-1, different-2, same-1, and same-2*. *Different-1* and *same-1* arrays contained icons from the same library, but never from icon set 2. One of the four types of arrays was designated as the S+, associated with reward. The other three, however, were not. On any given trial, only one type of array was displayed at a time. Pecks to the S+ displays delivered food regardless of whether the pigeons also pecked the S- stimuli. Each pigeon was therefore free to respond as it chose, with different patterns of responding to the four kinds of discriminative stimuli allocating attention in one of four ways: icon set alone, relation among the icons alone, both properties, or neither of the properties of the displays. In training trials, all four stimulus types were rewarded, while in discrimination trials, only one types of array was designated the S+. The authors argued, consistent with Goldstone and Barsalou (1998), that it may be more appropriate to view the task as falling somewhere on a perceptual-conceptual continuum. Goldstone and Barsalou (1998) stated that many phenomena we view as conceptual actually may be the result of perceptual processes becoming "less

bound to the perceptual specifics" (p. 256) of the stimuli. Therefore, the degree to which a process is conceptual is dependent upon the degree to which it is free of perceptual details (Wasserman, et al., 2002).

An interesting find from their previous study (Fagot et al., 2001) was a difference in the acquisition rates of the S/D concept for baboons. Criterion for *same* trials was attained more quickly than *different*. By varying the amount of variation in stimulus displays, Young and Wasserman (2002) examined properties that seem to make uniformity special. Entropy, the amount of variation within a display, is known to play an important role in the acquisition of abstract concepts. Pigeons as well as baboons and rhesus monkeys seem to have trouble detecting variety when a set consists of fewer than 8 items. By varying the number of distinct items within a *different* array, entropy is altered. For example, a 16-item array with all 16 distinct items (akin to those displays used in all previous Wasserman studies) has an entropy level of 4. However, a 16-item array with only 8 distinct items (each duplicated once) has an entropy of only 2.

In this study, Young and Wasserman (2002) evaluated the extent to which a display is classified as *different* based on *how* different it really is. Pigeon and adult human participants made S/D discriminations between displays while the arrays to be compared were not of equivalent entropy. Entropy varied from 0 (16 all identical) to 4 (16 all distinct) in increments of 0.5. When discrimination involved comparing a *different* array of entropy 4 to one of lesser entropy, S/D accuracy decreased with increased entropy. Along the entropy continuum, displays were discriminated asymmetrically: values at the lower end of the entropy scale were much more easily distinguished than those at the upper end of the scale. It was easier for both humans and pigeons (although

to a slightly lesser extent) to discriminate an entropy = 0 display from those with higher entropy than to discriminate an entropy = 4 display from those with lower entropy. When even a small amount of variability is introduced into a display, it becomes significantly more difficult to discriminate. These findings suggest that uniformity is saliently different from higher levels of entropy.

Critical Factors in Learning the S/D Distinction

In contrast to entropy-related studies in which the number of items *simultaneously* presented served as a critical factor, Katz, Wright, and Bachevalier (2002) identified *training stimulus set size* as crucial to S/D abstract concept learning by rhesus monkeys. An increased set size has the advantage of drawing attention from aspects of individual exemplars and placing emphasis on the relation between them. With small set sizes, individual features of objects may become the controlling cue, whereas in large stimulus sets individual features change frequently enough that stimulus relationships such as the S/D distinction are able to emerge as the basis for further discrimination (Katz et al., 2002). Further, with larger set sizes, issues of proactive interference are a more minimal factor. If the present stimulus is unique in the learning history of the animal, previous associations with that stimulus cannot be recalled. Katz and colleagues. (2002) found that an increase in training set size is associated with a decrease in the number of trials to criterion for rhesus monkeys. In addition, with larger set sizes, higher levels of accuracy were achieved.

After being trained on a two-choice S/D discrimination task using a set size of only 8 photographic stimuli, animals began a transfer test with novel sets of stimuli. New sets contained between 8 and 128 unique stimuli. All animals completed transfer tests

stimuli sets of several different sizes. As set size increased, the number of trials to

criterion decreased. The more exemplars contained in the stimuli set, the more quickly

the same/different distinction was made. However, for sets above 32 exemplars, animals

all performed significantly above chance after only approximately 100 trials. Rather than

being an "all-or-nothing" phenomenon, abstract concept learning may rely heavily on

larger set sizes. Abstract conceptual learning with significantly smaller stimulus sets may

never emerge, whereas an abundance of exemplars help animals to break free from a

predisposition to item-specific learning (Katz et al., 2002).

Capuchin monkeys were presented with a task similar to that completed by rhesus

monkeys in previous experiments (Katz et al., 2002). As set size increased, transfer

performance also increased. However, capuchin monkeys learned the S/D task and

abstract concepts much more rapidly than rhesus monkeys trained in the same procedure.

Capuchins tended to learn the task in one-fifth the amount of transfer sessions as rhesus

monkeys.

Whereas Fagot and colleagues (2001) have demonstrated that at least two

baboons would have marked difficulty on a relational match-to-sample (RMTS) task,

rhesus monkeys at the Language Research Center (LRC) with extensive testing

experience were presented with the task in order to verify these results with trial-unique

stimuli (Flemming et al., 2005). In an experiment motivated by the views of Washburn,

Thomspon, and Oden (1997), pairs of images were used as stimuli in a matching-to-

sample task. One stimulus pair (composed either of two identical or two different trial-

unique randomly drawn images) served as the sample. After contact was made with this

object (via a cursor controlled by a joystick), choice pairs were presented on each side of

the bottom half of the screen. One choice pair was composed of two identical images (but different from any of those present in the sample pair) and the other was made of two images that differed physically from each other. Monkeys were required to choose the pair which matched the same relation (either *same* or *different*) of the sample pair. After verifying that monkeys fail on such a task, several other steps were taken to determine why they failed and whether they could be trained to succeed in such a task.

A series of experiments was designed beginning with the most basic of discriminations: a two-choice paradigm (Flemming et al., 2005). Throughout the course of the experiments, several conclusions could be made about the acquisition of concepts by rhesus monkeys: implicitly, their concepts for *same* and *different* may be better described as *uniformity* and *chaos*, respectively, increasing the entropy of a display acts as a means to learn said concepts, and discriminative cues are integral for proficient success. In Experiments 1 and 2, we found that pairs of images may present a unique problem for the monkeys. In a simple two-choice discrimination paradigm, monkeys completed over 10,000 trials with chance levels of accuracy. Paired images were spread farther apart by space and separated with lines (see Figure 2). However, side biases revealed that all five monkeys had no particular strategy with which to solve the task. From this, we concluded that the monkeys may not have perceived the pairs as such. That is, they were seeing the pair of images as one stimulus "bunch" rather than a pair of two identical or nonidentical images.

In Experiment 3, groups of eight identical and nonidentical stimuli were presented in place of pairs (see Figure 3). Monkeys quickly learned their assigned S+.
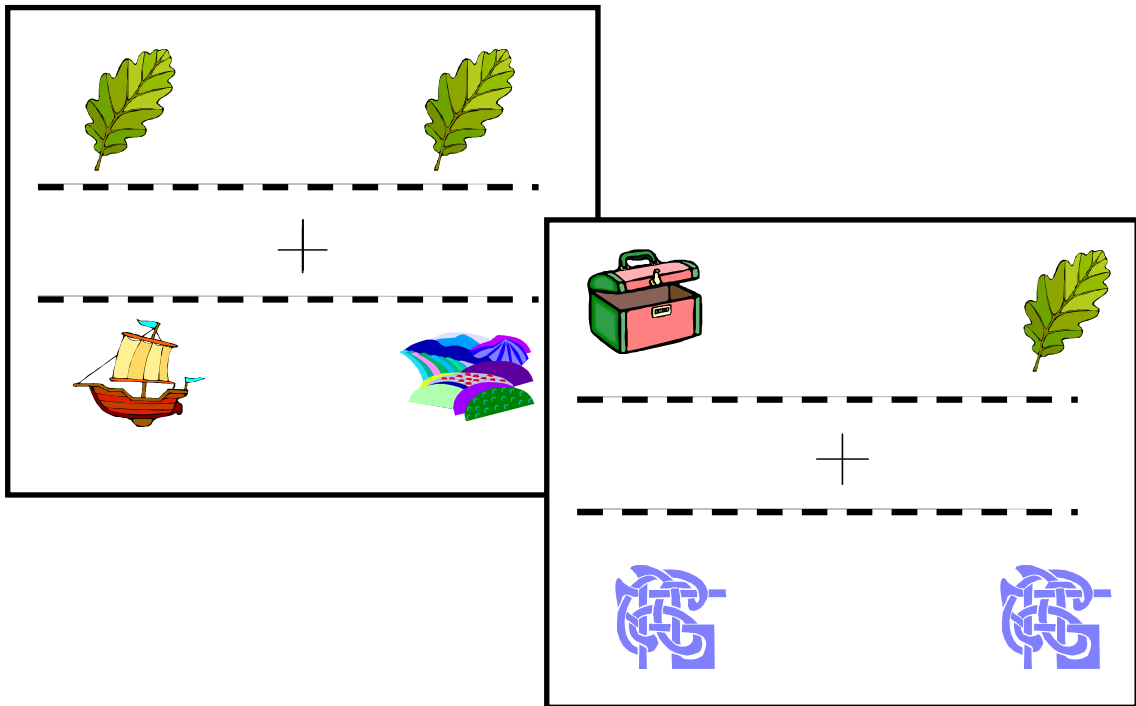
Figure 2. Paired-Image Displays Presented to Rhesus Monkeys in Experiment 2 of Flemming, Beran, and Washburn, 2005.
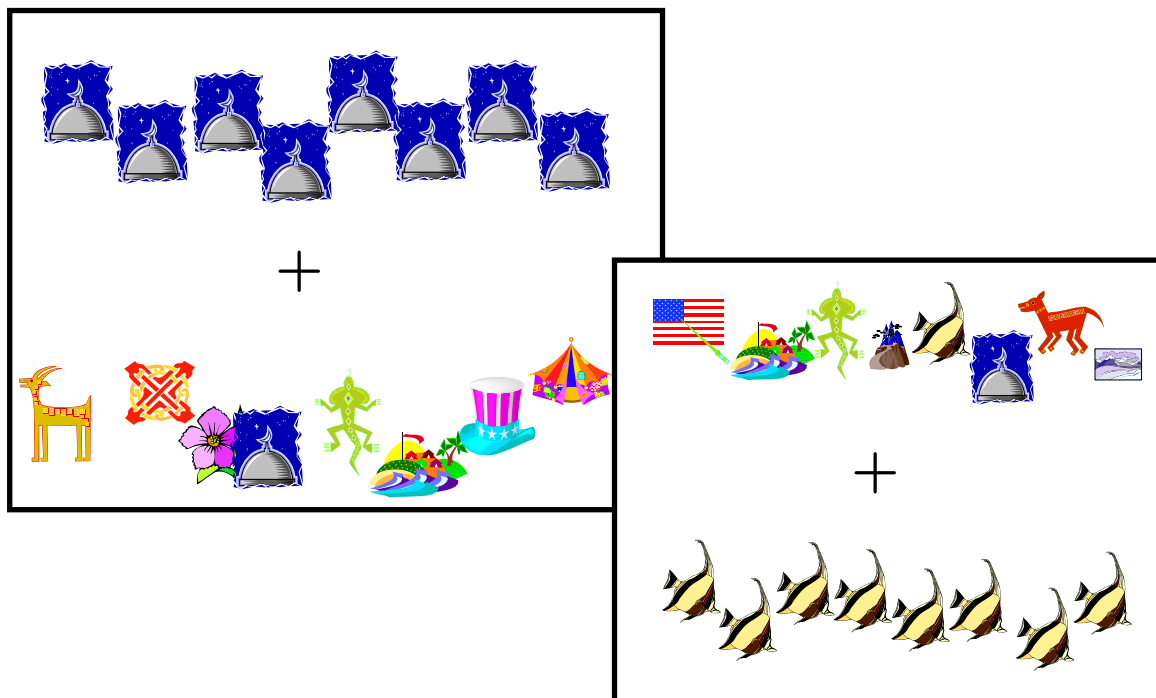
Figure 3. <u>Entropy-Infused Displays Presented to Rhesus Monkeys in Experiment 3 of Flemming, Beran, and Washburn, 2005.</u>

However, when the S+ switched from *same* to *different* (or from *different* to *same*), all monkeys perseverated on their initially rewarded stimulus for thousands of trials.

To give the monkeys a better indication of the S+, we introduced background color as a discriminative cue in Experiment 4. With the entropy of the displays increased (set size = 8), monkeys successfully discriminated between rows of identical and nonidentical images when accompanied by an S+ specific background color. When *same* was the designated S+, the background color was pink. A black background served as an indicator for an S+ of *different*. We hypothesized that these background colors functioned as referential labels for the rewarded stimulus. Gradually, the number of items in the displays was reduced to a pair, which now proved no more difficult than sets of 8. In contrast to Fagot, Wasserman, and Young's (2001) work with baboons and Wasserman, Frank, and Young's (2002) work with pigeons, we found that rhesus monkeys are capable of accurately making the distinction between pairs of identical and nonidentical images.

Abstract Concepts and Language

Premack (1976, 1983, 1986) suggested that, at least in the case of abstract relations, acquisition of conceptual knowledge depends upon language. Therefore, it should come as no surprise that while language-naïve nonhuman animals show varying degrees of perception-based conceptual knowledge, they lack the capacity to understand abstract concepts such as *same* and *different* to the same extent as humans understand them (Premack, 1983, 1986; Thompson, 1995; Thompson & Oden, 2000).

Language has further been implicated as the main mechanism responsible for judgment of relations-between-relations—that is, second-order relations (Premack, 1976, 1983; Premack & Premack, 2003; Thompson & Oden, 2000).

Before Wasserman and colleagues demonstrated that pigeons and some baboons may have this ability (i.e., Fagot et al., 2001; Wasserman et al., 2002), only adult humans and language-trained chimpanzees had been shown to display abstract relational matching skills. Thompson, Oden, and Boysen (1997) revisited this accusation by presenting language-naïve chimpanzees with a conceptual matching-to-sample task. After being familiarized with a physical match-to-sample task, five adult chimpanzees viewed paired random junk objects as sample and choice stimuli. As in several previous tasks (i.e., Fagot et al., 2001), the goal was to indicate the choice pair that conveyed the same relation between the objects as the sample pair.

Four of five chimpanzees spontaneously judged the conceptual equivalence of relations-between-relations. The fifth chimpanzee differed in his learning history; he was naïve with respect to numeric problem solving and symbolic token training in addition to language training. Therefore, it seems that these tokens may have a functional role in the acquisition of abstract concepts. Thompson and Oden (1996) suggested that the critical role of the token is to provide an animal with a concrete icon for encoding a propositional representation that is otherwise abstract. In the context of abstract relational matching-to-sample, the token may "objectify" a relationship or have the retrieval function of a word (Thompson et al., 1997). Thompson, Oden, and Boysen (1997) also suggested that conceptual-relational matching is akin to covert symbol matching.

Thompson and colleagues (1997) hypothesized that the judgment of relations-between-relations is made possible by the animal's representational capacity to re-encode abstract relations into iconically equivalent symbols. It should follow then, that such symbolic training produces a system for universal computation.

By comparing directly three distinct primate species, differences in performance may reflect evolutionary change as a function of relatedness to a common primate ancestor. Because analogical reasoning cannot be reduced to stimulus-response (S-R) learning, it is reasonable to expect that, like emergent behaviors, we should note a qualitative shift in relation to cranial capacity (Rumbaugh, 2002). Because rhesus monkeys are more distantly related to humans than are chimpanzees, similarities in analogical reasoning skills may suggest that the phenomenon was in existence for a large part of primate evolutionary history. Chimpanzees diverged from the common primate ancestor more recently than rhesus monkeys, with respect to humans. Therefore, behaviors observed only in chimpanzees and humans, but not rhesus monkeys, may be a more recent development in cognitive evolution.

The purpose of the current project was to determine the role of meaning in the analogical reasoning abilities of three different primate species. Such comparisons outline further our understanding of the cognitive capacities of various species and further our knowledge about concepts and mechanisms of concepts learning in general. According to Thompson and colleagues (1997), the capacity to re-code abstract relations (such as same-different) is solely responsible for success in conceptual relational matching. Such a recoding would require animals to have an explicit symbol or token consistently associated with the abstract relations themselves. Perhaps, rather than a strict language

system, meaningfulness of stimuli plays the pivotal role in the acquisition of abstract concepts. According to this hypothesis, as long as meaning can be attributed to individual stimuli, nonhuman animals will begin to recognize pairs as matched or nonmatched in meaning, rather than perception.

<div align="center">Method</div>

Participants

Three groups of participants were tested on the analogical reasoning task: adult humans, rhesus monkeys (*Macaca mulatta*), and chimpanzees (*Pan troglodytes*).

Eighty-two undergraduates (67 females) were recruited from Georgia State University's psychology research pool with half assigned to each stimulus condition (meaningful and nonmeaningful stimuli) in the relational match-to-sample task. The mean age of the participants was 20 years and 60% were minority students. All participants completed an informed consent form and received debriefing instructions upon completion of the task.

Four chimpanzees (18 to 34 years of age) housed at Georgia State University's Language Research Center (LRC) were also tested. The chimpanzees had previously participated in experiments involving the simple match-to-sample paradigm with joysticks (Beran & Washburn, 2002), but were naïve to the specific testing procedures involving the S/D concepts. Individuals were randomly assigned to receive one of two conditions (meaningful and nonmeaningful lexigram stimuli) first, followed by completion of the remaining condition. The chimpanzees were not food or water deprived. Individuals worked at mobile testing systems at their home cages for designated 1-hour sessions each day.

Five male rhesus monkeys (10 to 20 years of age) housed at the LRC also completed the relational match-to-sample task. Like the chimpanzees, the monkeys had been previously trained on simple physical match-to-sample tasks (MTS) with joysticks where correct responses were exact physical matches to target stimuli (Washburn et al., 1992; Washburn & Astur, 1998). Individuals were randomly assigned to receive one of two conditions (meaningful and nonmeaningful stimuli) first, followed by completion of the remaining condition. The monkeys were not food or water deprived for purposes of testing, and allowed to work *ad libitum* throughout the day in their home cages.

All monkeys began this study with previous experience on other tasks involving the same-different paradigm. In the Flemming, Beran, and Washburn (2005) study, monkeys discriminated between arrays composed of identical or nonidentical clipart images. In the presence of a discriminative color cue, monkeys successfully discriminated same from different displays of 8, 6, 4, and 2 items. In addition, monkeys completed, but failed RMTS tasks similar to those in the current study, but composed only of clipart images.

Apparatus

The LRC's Computerized Test System (LRC-CTS) consists of an IBM-compatible desktop personal computer (Washburn et al., 1992). This same apparatus was used throughout all parts of the project, with slight variations for each species. Undergraduate students at Georgia State University were tested at a desktop computer using a hand-held joystick. Each nonhuman animal had access to its own testing station. During tasks, monkeys controlled a cursor on a 17-inch SVGA monitor via a vertically-mounted joystick. The monitor was positioned approximately 15 cm from the home cage

behind a transparent Lexan plate. Chimpanzees controlled a horizontally-mounted joystick within a port attached to their home cages; stimuli were presented on a monitor approximately 1m outside of the home cage on a mobile cart.

Speakers provided auditory feedback for all tasks, including a low buzzing sound for incorrect choices and an increasing crescendo sound for correct choices. These sounds have been paired with these outcomes on many previous tasks. For the current tasks, the increasing crescendo sound was always accompanied by the dispensing of a 94-mg banana-flavored pellet to rhesus monkeys and small portions of fruit or 1-g pellet to the chimpanzees.

Design and Procedure

To investigate the role of meaning in the acquisition of abstract relational concepts in the proposed study, the three different species completed comparable tasks. Human adults, chimpanzees, and rhesus monkeys completed two relational match-to-sample (RMTS) tasks that differ on one dimension. In one condition, the stimuli conveyed no meaning to the participant, whereas in the other condition, stimuli carried discreet, specific meanings. Participants completed both tasks in a randomized design, with some individuals receiving the meaningful condition first and others receiving the nonmeaningful condition first.

Task

In the RMTS task, stimuli were presented in pairs with one sample pair and two choice pairs. At the initiation of a trial, one pair of stimuli (either two identical or nonidentical objects) was centered at the top of the computer screen; stimuli were approximately 5 cm x 3 cm. Participants were required to contact this sample pair with

the stimuli in order for the choice pairs to appear. Once contact with the sample pair was made, the joystick cursor was re-centered on the screen, and the choice pairs appeared in the bottom half of the screen on the left and right sides. One choice pair contained two identical items, whereas the other contained two physically nonidentical items. Importantly, no stimulus in the choice pairs was ever physically identical to stimuli in the sample pair and was randomly assigned to position. The task, then, for example, was to match AA with BB (and not CD) and to match EF with GH (and not JJ). Similar testing paradigms have been frequently utilized (Premack, 1976; Thompson & Oden, 2000; Flemming et al., 2005). Successful performance of the conceptual-relation matching task required that the participants judge one relation to be the same or different from *another relation*.

To make a response, the cursor was moved either left or right toward the choice pairs. Once contact was made with a choice pair, a feedback sound was played (an increasing tone if the choice was correct or a buzzing if incorrect) followed by rewards on correct trials for nonhuman animals and a short intertrial interval (ITI) until the sample pair for the next trial appears. For correct choices, rewards were automatically dispensed to the animals accompanied by a 2-s ITI. When choices were incorrect, no food reward was dispensed and longer ITIs were imposed (15 s for rhesus monkeys and 5 s for chimpanzees). Humans, however, received instruction as to the meaning of the feedback sounds. This was the *only* instruction they received. In addition, humans received no food reward and 2-s ITIs for both correct and incorrect choices. After the ITI, the next trial was automatically initiated and the next sample pair appears at top of the screen.

Stimuli were selected from one of two separate conditions for each participant or animal: meaningful or nonmeaningful. In the meaningful condition, stimuli were discreet symbols that had meaning specific to the species.

Stimuli

Humans. Meaningful stimuli were composed of 3- to 7-letter words that referred to concrete objects such as foods and places. Nonmeaningful stimuli were composed of 3- to 7-letter strings of illegal nonwords generated by the program WordGen (Duyck, Desment, Verbeke & Brysbaert, 2004). Illegal nonwords were chosen as appropriate nonmeaningful stimuli because they carry with them no inherent referential value, and cannot be recoded into sensible phonemes in the English language. Both sets of stimuli appeared as white letters inside a black rectangle. Figure 4 portrays example displays presented to human participants.

Chimpanzees. Lexigrams are symbolic characters that convey special meaning of real-world objects to the chimpanzees. Through specialized training, the animals learned to communicate with the researchers and one another about certain foods, places, people, and activities (Rumbaugh, 1977; Savage-Rumbaugh, Fields, & Taglialatela, 2001; Savage-Rumbaugh et al., 1980; Rumbaugh & Washburn, 2003). The LRC chimpanzees have been shown to retain the meaning of these symbols for more than 20 years (Beran, Pate, Richardson, & Rumbaugh, 2000). In addition, the animals sort lexigrams into labeled groups more accurately than real-world objects and photographs (Savage-Rumbaugh et al., 1980).
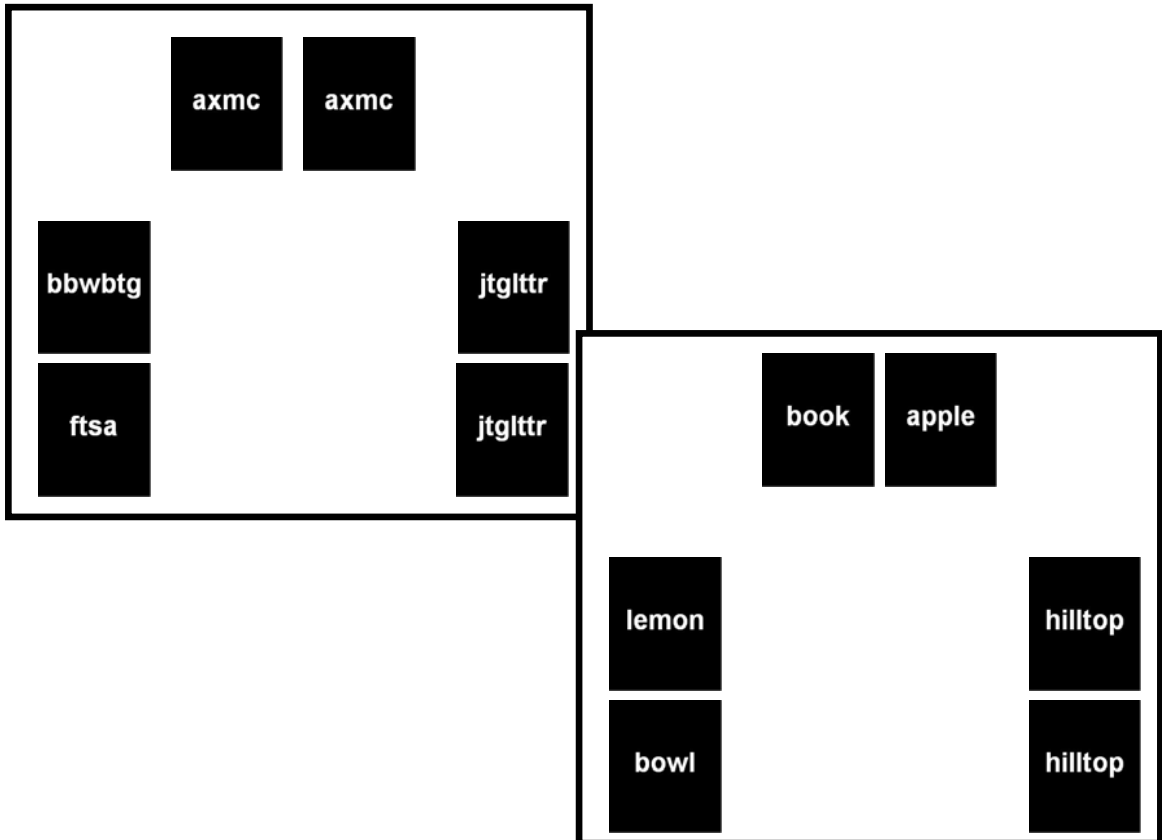
Figure 4. Example Displays Presented to Human Participants

Only lexigrams with which the chimpanzees have had extensive experience and have been shown to retain meaning for the animals were presented during the meaningful condition. These lexigrams may function more like whole words for the animals carrying with them a specialized meaning (Rumbaugh & Washburn, 2003). Exemplars from an unknown, never before seen subset of lexigrams were used as nonmeaningful for chimpanzees (see Figure 5).

For the first two sessions (20 to 30 trials) animals were hand-fed with preferred fruits after the selection of correct choices. Because animals were distracted by experimenter presence, oftentimes not attending to the computer screen during trials, the automatic pellet dispenser was utilized throughout subsequent testing sessions.

Monkeys. Although the rhesus monkeys had seen lexigram stimuli in previous studies, these animals had never been trained on lexigrams meanings. There is sufficient evidence however that monkeys can at least use symbols as labels for certain concepts. In tasks presented to the rhesus monkeys at the LRC, Arabic numerals have been paired with specific numbers of pellets, possibly allowing the animals to associate these numeric symbols with quantity information (Rumbaugh & Washburn, 1991; Washburn, 1994).

In addition to simple quantity information that numerals may convey, Arabic numerals have also produced Stroop-like effects in rhesus monkeys at the LRC. Because meaning of the numerals interferes with judgments about amount, numerals mean amounts, and thus can be said to have symbolic representation for the monkeys (Washburn, 1994).
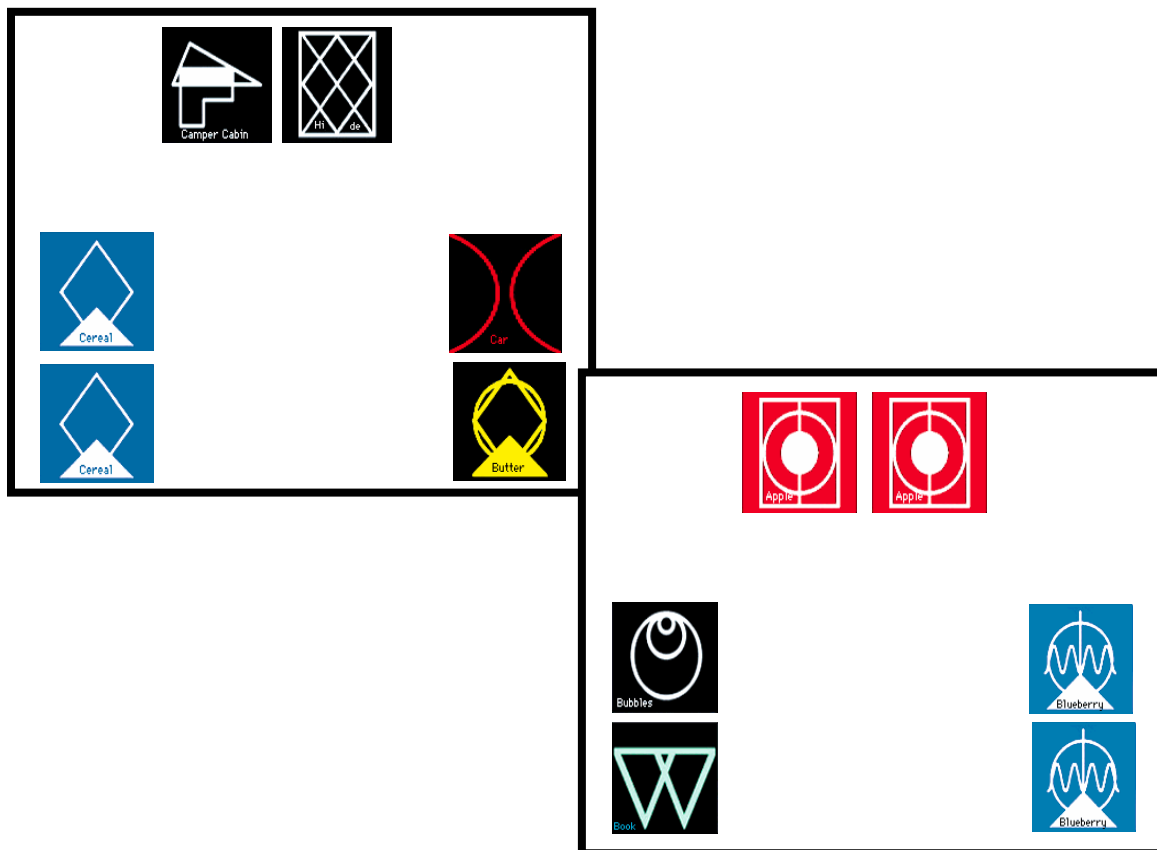
Figure 5. Example Displays Presented to Chimpanzees

Adding to the repertoire of what Arabic numerals mean to the rhesus monkeys at the LRC is the predictive power of the numeric symbols. Harris & Washburn (2005) presented the monkeys with series of reinforced and nonreinforced maze trials in which Arabic numerals indicated the number of reinforced mazes that must be completed before a nonreinforced maze trial was presented. Monkeys developed a pattern of responding slower on nonreinforced trials than the preceding reinforced trial; they used the Arabic numeral as a cue to the number of reinforced maze trials that would occur in a series (Harris & Washburn, 2005). These previously learned Arabic numerals were utilized in paired stimuli in the meaningful condition (see Figure 6).

In the nonmeaningful condition, stimuli consisted of basic Latin alphabet letters. Special attention was given so that letters used in previous tasks for specific choice responses were not included in the subset of letters used in this task. Choice of letters and numerals controlled for perceptual qualities of the stimuli by balancing their basic physical properties.

Results

Humans

Accuracy-by-condition was assessed by comparing the average performance for the entire 100-trial block between the two groups of participants (*meaningful/ nonmeaningful*). Participants in the meaningful condition completed the task with an average accuracy of 93.86 out of 100 trials (SE = 1.73), significantly higher than participants in the nonmeaningful condition (M = 87.48, SE = 2.47); $t(80) = 2.13$, $p < .05$, $d = 6.38$.
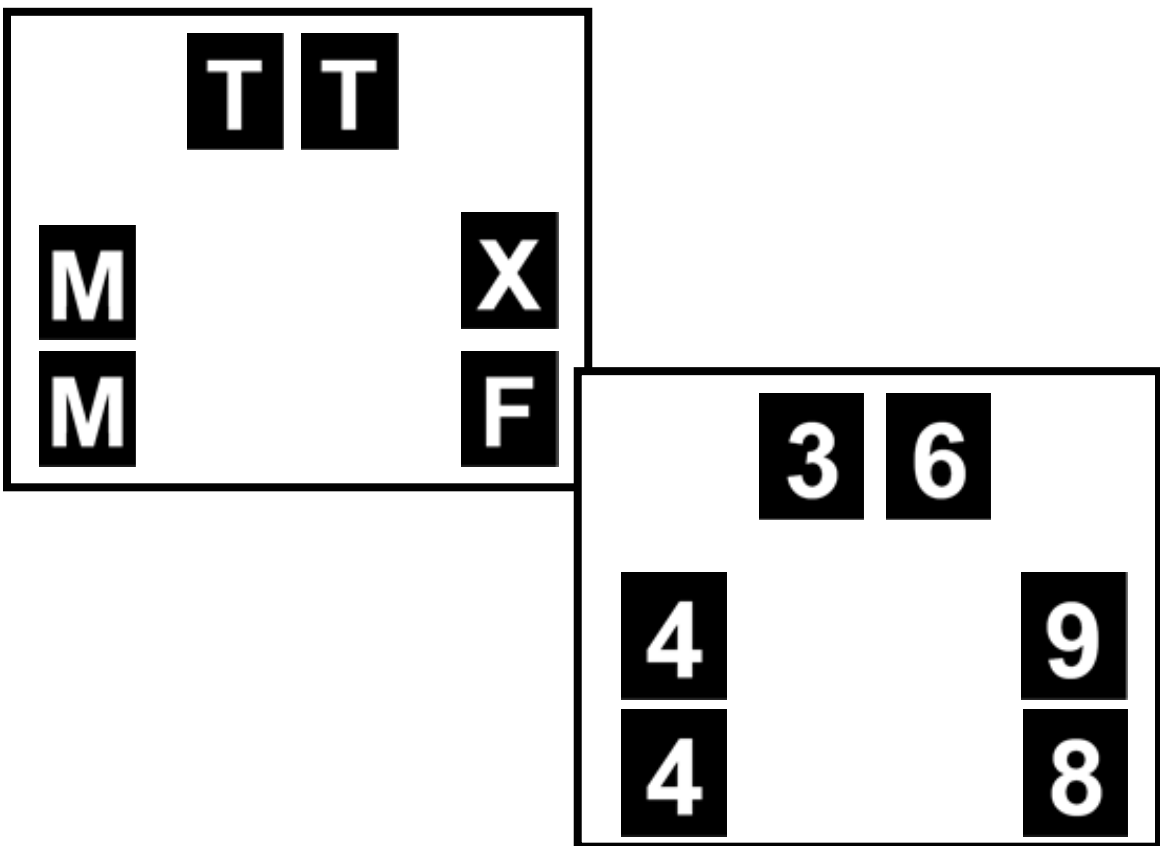
Figure 6. Example Displays Presented to Rhesus Monkeys

Within a 100-trial testing session of either the meaningful or nonmeaningful stimuli, 76 of 82 participants met the criterion (correct responses to 8 out of the previous 10 trials attempted) for sufficient acquisition of an analogical rule. Two participants who received the meaningful condition never met a criterial level of accuracy; four participants assigned to the nonmeaningful condition did not perform at levels significantly above chance (50%) after 100 trials. These six participants were removed from analysis for trials-to-criterion and response time.

Because accuracy was generally high for participants in both conditions, the number of trials-to-criterion was assessed to determine how long participants failed before realizing the analogical rule. Trials-to-criterion for each participant was calculated by summing the total number of trials until eight out of the previous ten trials attempted were completed correctly. Participants in the meaningful condition met criterion on average within 14.35 trials (SE = 1.38), fewer than the number of trials required for the participants in the nonmeaningful condition (M = 19.08; SE = 2.60); however, this difference was not statistically significant, $t(74) = 1.65$, $p = .10$, $d = 4.73$ (see Figure 7).

Response time (time from the appearance of the choice pairs to the selection of one choice) provides another measure of learning by assessing how long participants process the stimuli before making a choice. A mixed-design ANOVA revealed a significant within-subjects main effect (*RT before/RT after*), but no significant interaction with the between-subjects factor (*meaningful/nonmeaningful*). In both conditions, response time was significantly shorter after criterion was met, $F(1,74) = 61.64$, $p < .01$ (see Figure 8).
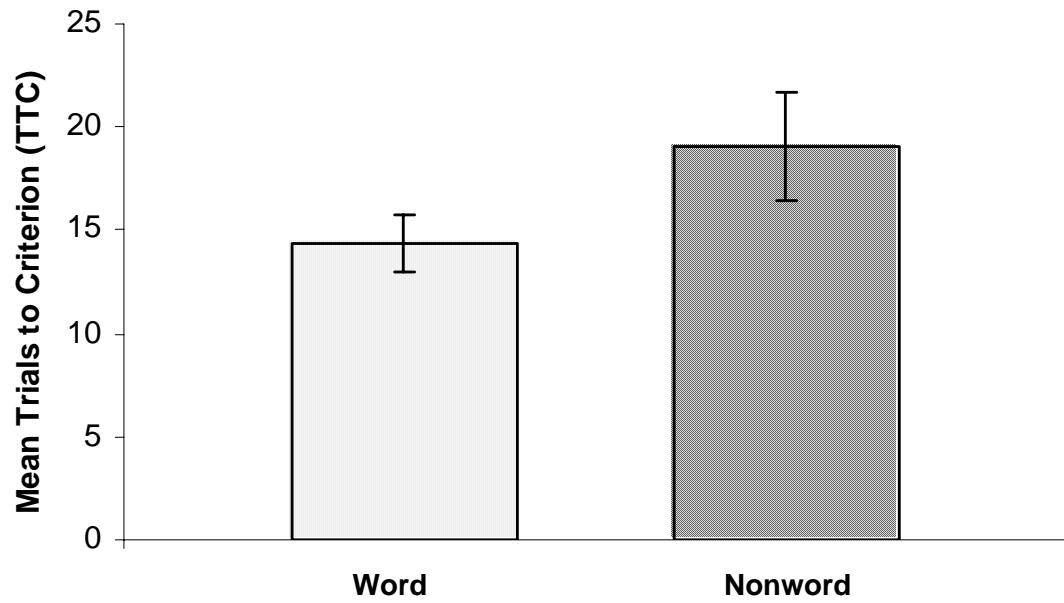
Figure 7. <u>Number of Trials to Reach Criterion for Human Participants (error bars</u>
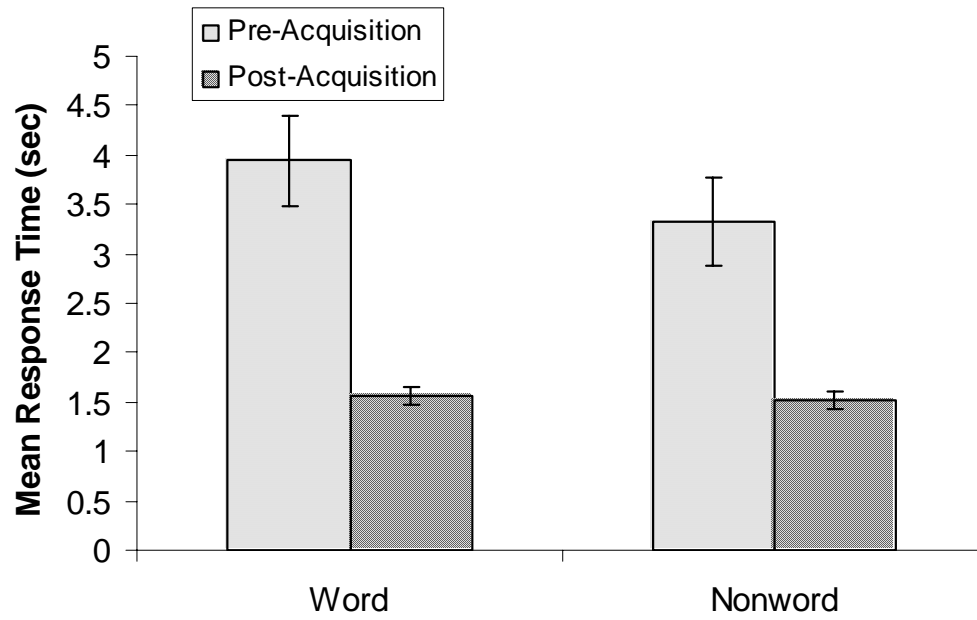
<u>represent standard error)</u>

Figure 8. Response Times Pre- and Post-Acquisition for Human Participants (error bars represent standard error)

Response times both before (*RT before*) and after (*RT after*) meeting criterion were statistically equivalent for both conditions, $F(1,74) = 0.88$, p = .35.

Chimpanzees

There were individual differences in performance by the chimpanzees on the RMTS task. Table 1 presents performance summaries for the chimpanzees. Two chimpanzees (Lana and Mercury) never performed significantly above chance in any condition—meaningful (lexigrams or numerals) or nonmeaningful—after at least 110 trials in each condition.

In several testing sessions, these animals developed positional biases, choosing either the left or right stimulus exclusively for a significant number of trials. Mercury (who received the nonmeaningful condition first) contacted the left pair on 81% of all trials in the numeral condition, significantly more often than the right pair $z = 6.48$, $p <$ .01. In meaningful and nonmeaningful conditions, however, Mercury developed strong right-side biases: 76%, $z = 7.16$, $p < .01$ for meaningful trials; 66%, $z = 3.73$, $p < .01$ for nonmeaningful trials. To illustrate that performance did not improve across trials, Figure 9 presents Mercury's cumulative percentage of correct responses per trial block in the three conditions.

Performance was found to be symmetric in all conditions for Mercury. That is, accuracy on *same* trials and *different* trials was statistically equivalent across conditions "meaningful" lexigrams $\chi^2 (1, 191) = .09$, $p > .05$; nonmeaningful lexigrams $\chi^2 (1, 133) =$ .81, $p > .05$; meaningful (numerals $\chi^2 (1, 110) = .003$, $p > .05$).

Table 1.

Performance summaries of chimpanzees

Subject	Condition	# of Trials Completed	Percent Correct
Lana	Meaningful (lexigrams)	238	52
	Nonmeaningful	166	54
	Meaningful (numerals)	342	52
Mercury	Meaningful (lexigrams)	191	49
	Nonmeaningful	133	56
	Meaningful (numerals)	110	48
Panzee	Meaningful (lexigrams)	202	66*
	Nonmeaningful	186	54
	Meaningful (numerals)	286	44
Sherman	Meaningful (lexigrams)	227	55
	Nonmeaningful	263	65*
	Meaningful (numerals)	166	54

* $p < 0.05$

Note. Number of trials for each individual and condition varies as a result of how many trials each chimpanzee was able to complete in a given testing session. These data represent approximately 12 sessions per animal over the span of 21 weeks of testing.
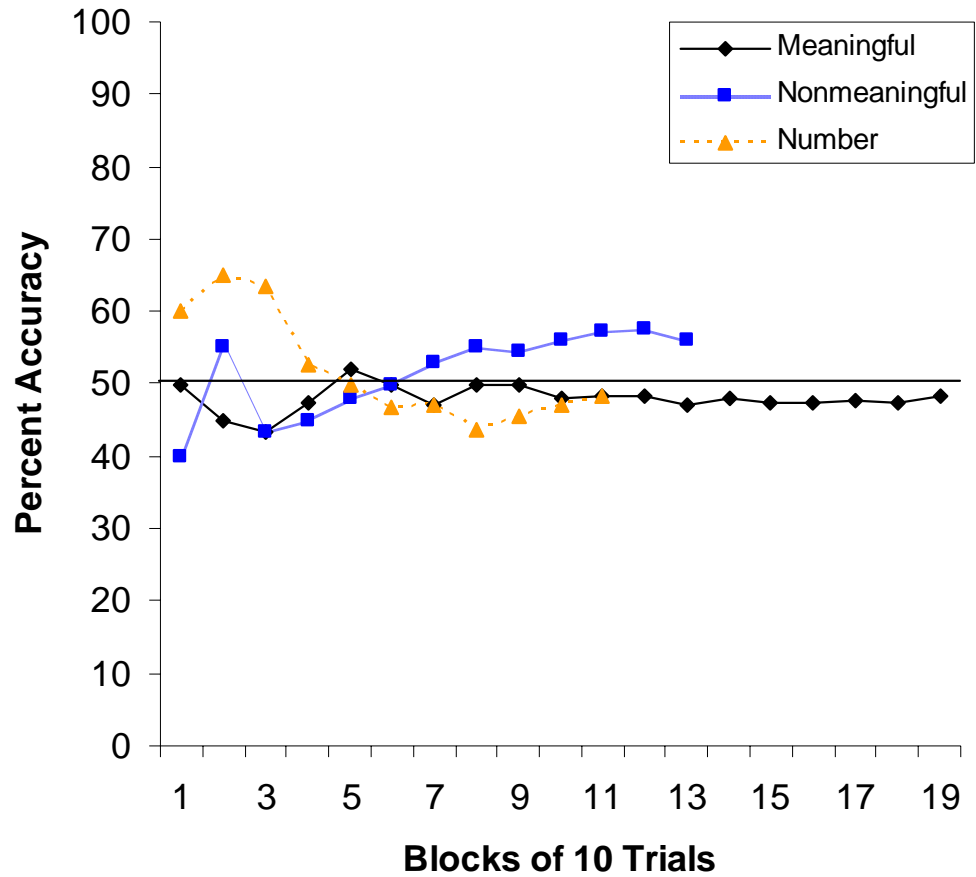
Figure 9. Percent Correct by 10-Trial Block for Mercury. The solid horizontal line at 50% represents chance performance.

Lana (who received the meaningful-lexigrams condition first) also developed a similar side bias with the meaningful lexigrams, selecting the right stimulus pair in 68% of all trials, $z = 5.70$, $p < .01$. In the nonmeaningful condition, however, Lana developed a left side bias (68%) in the first 100 trials ($z = 3.60$, $p < .01$), and then shifted strategies to a right bias (67%) in the remaining 66 trials ($z = 2.71$, $p < .01$). To illustrate that performance did not improve across trials, Figure 10 presents Lana's accuracy per trial block in the three conditions.

Performance was found to be symmetric in all conditions for Lana. Accuracy on *same* trials and *different* trials was statistically equivalent for meaningful lexigrams χ^2 (1, 238) = .96, $p > .05$; nonmeaningful lexigrams χ^2 (1, 166) = .24, $p > .05$; and for meaningful numerals χ^2 (1, 342) = 1.78, $p > .05$).

Panzee completed trials with meaningful stimuli before receiving the nonmeaningful and numeral conditions. She performed at levels above chance in the meaningful condition completing 66% of 202 trials correctly, $z = 4.64$, $p < .01$, but failed to perform above chance levels in the nonmeaningful and numeral conditions. Figure 11 presents Panzee's percentage of correct responses as a function of trial block and the three conditions.

Distribution of accuracy varied by condition for Panzee. Accuracy on *same* trials and *different* trials was statistically equivalent for meaningful conditions: lexigrams χ^2 (1, 202) = 2.06, $p > .05$; numerals χ^2 (1, 286) = 2.04, $p > .05$. In the nonmeaningful condition (which yielded overall chance performance), accuracy was significantly higher on *different* trials (65%) than on *same* trials (49%), χ^2 (1, 186) = 4.61, $p < .05$.
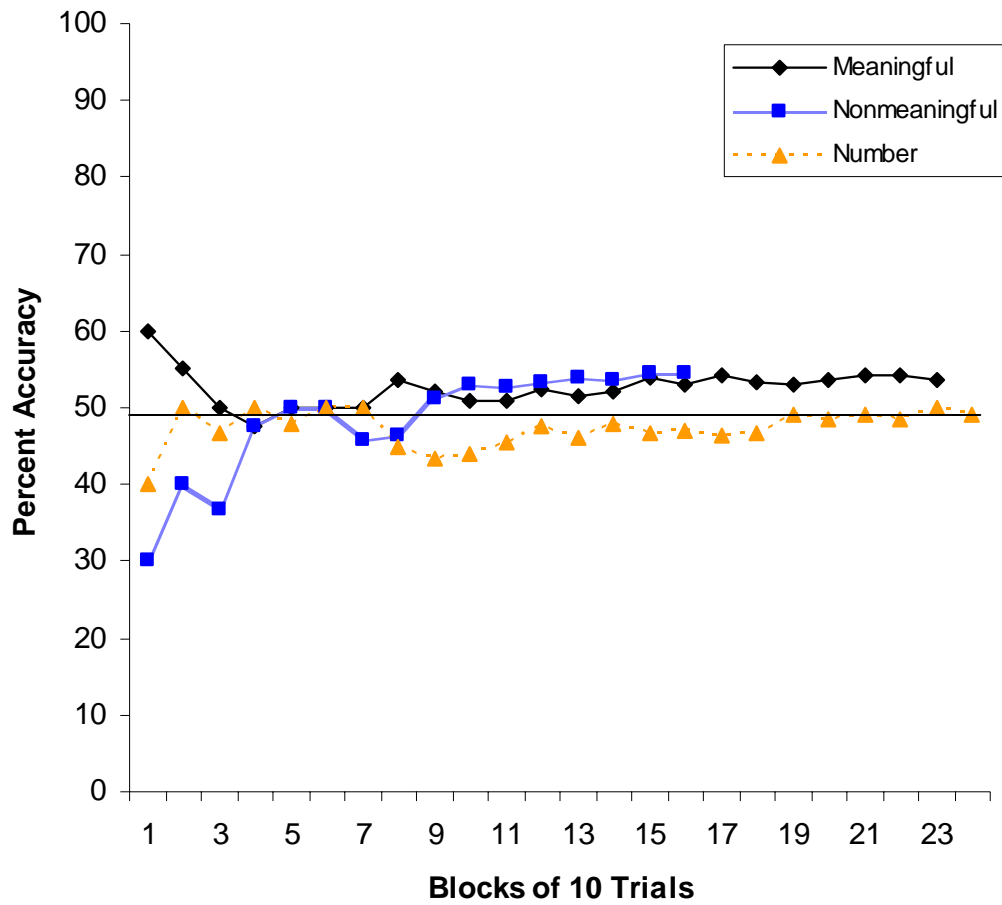
Figure 10. <u>Percent Correct by 10-Trial Block for Lana</u> The solid horizontal line at 50%

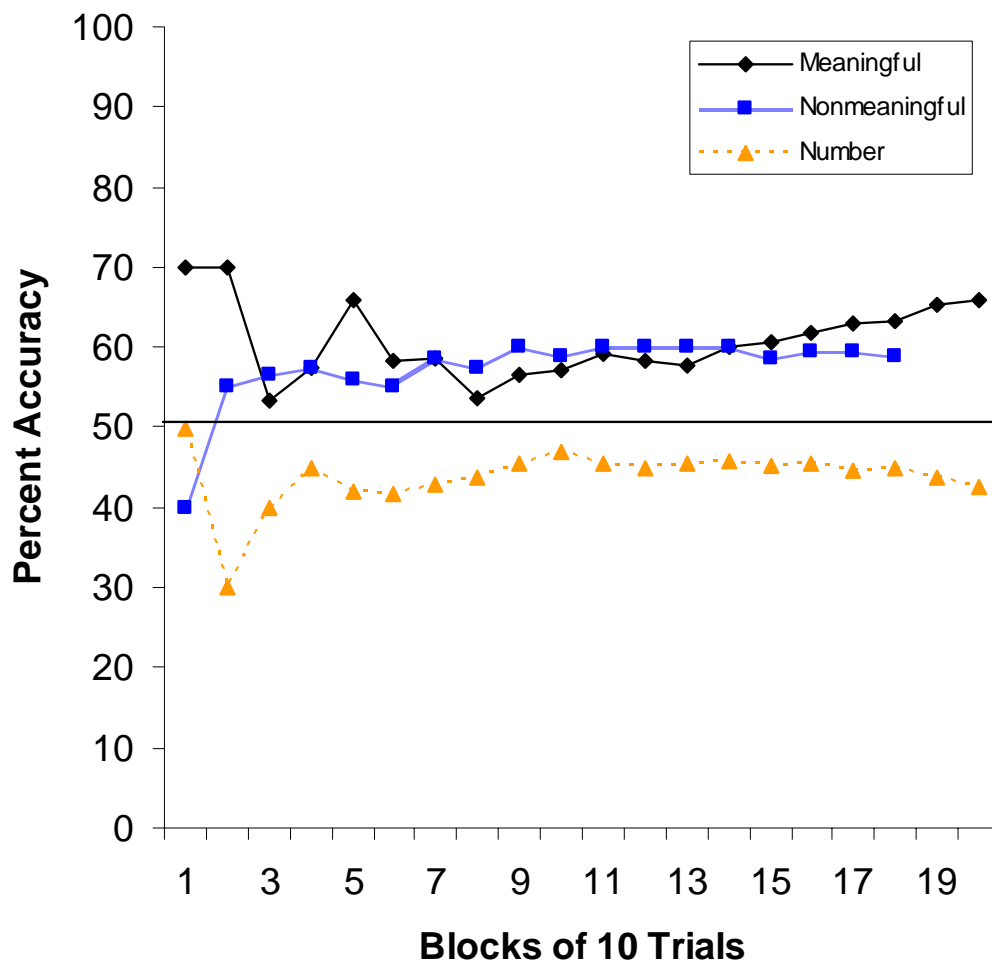represents chance performance

Figure 11. <u>Percent Correct by 10-Trial Block for Panzee.</u> The solid horizontal line at 50% represents chance performance

Sherman completed trials with nonmeaningful stimuli before receiving the meaningful and number conditions. He performed at levels above chance in the nonmeaningful condition, completing 65.00% of 263 trials correctly, $z = 4.99$, $p < 0.01$; but he failed to reach significance in the meaningful conditions. Figure 12 presents Sherman's percentage of correct responses as a function of trial block and the three conditions.

Distribution of accuracy varied by condition for Sherman. Accuracy on *same* trials and *different* trials was statistically equivalent for meaningful conditions: lexigrams χ^2 (1, 227) = 0.48, $p > .05$; numerals χ^2 (1, 166) = 1.45, $p > .05$. In the nonmeaningful condition accuracy was significantly higher on *different* trials (80.6%) than on *same* trials (50%), χ^2 (1, 263) = 27.18, $p > .05$.

For each chimpanzee, accuracy levels were consistent across all trials. No significant increases in performance across time were observed; no performance plateau was reached until the last few trials completed. In instances where performance failed to reach significance, levels of accuracy remained around chance throughout the testing period. For those instances where chimpanzees did perform above chance levels, acquisition was spontaneous, rather than resembling gradual learning across trials.

Rhesus Monkeys

Order of presentation for the two conditions, meaningful and nonmeaningful was counterbalanced with three monkeys receiving nonmeaningful stimuli (letters) first and the other two monkeys receiving meaningful stimuli (numerals) first.
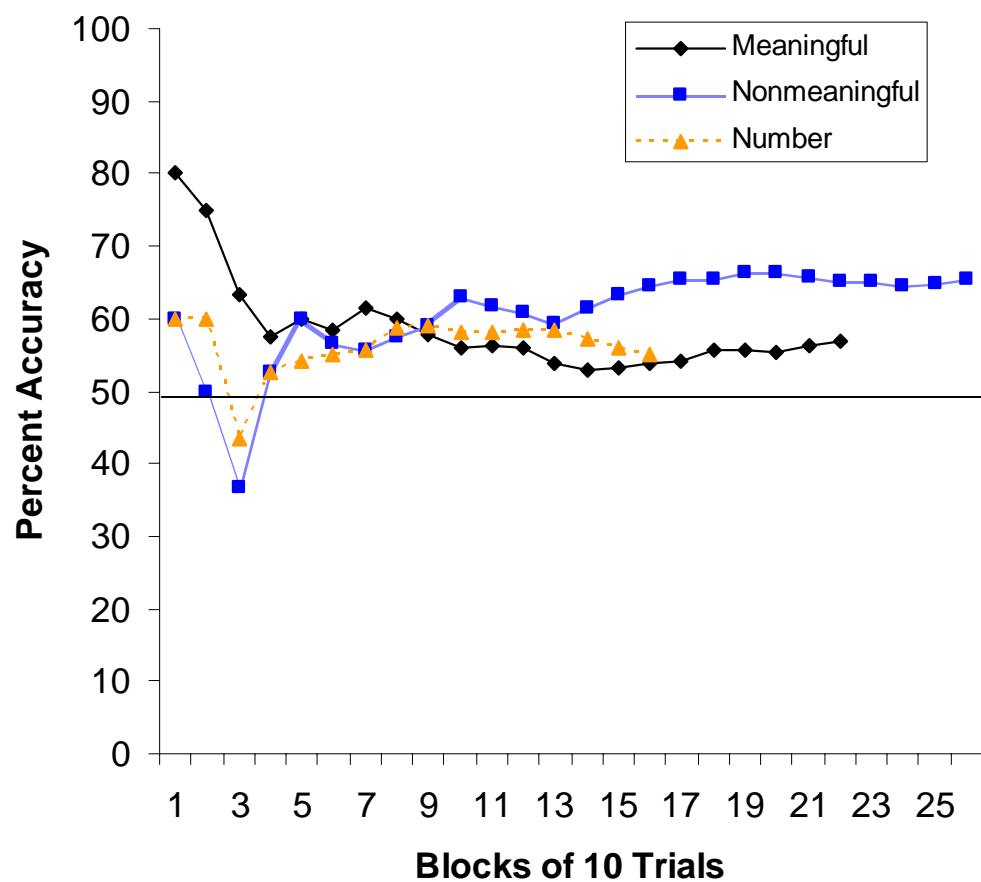
Figure 12. <u>Percent Correct by 10-Trial Block for Sherman</u> The solid horizontal line at

50% represents chance performance

The monkeys completed an average of 1,937 trials in each condition, but no monkey achieved levels of performance significantly above chance on either condition, regardless of which was presented first. Performance summaries of the rhesus monkeys are shown in Table 2, with illustrative accuracy levels displayed in Figure 13 for Hank. The trends of the learning curves presented in Figure 13 are representative of the performance of all other monkeys in this study. Performance failed to improve across trials, and position biases emerged for every animal and condition.

If performance gradually improved across trials, eventually approaching significance, we could conclude that the animals slowly learned an analogical rule through trial-and-error. This is not the case, as no animal ever reached a level of performance significantly different from chance. To illustrate that there is no evidence that performance did not improve across trials, blocks of 100 trials were analyzed independently for Hank for the first 1,000 trials in both conditions.

Percent correct was not significantly different from chance in any trial block except for trials 601-700 (60% correct; $z = 2.0$, $p < 0.05$) in the nonmeaningful condition. Subsequent trial blocks rebounded to near chance performance. Table 3 summarizes these data.

Table 2.

Performance Summaries of the Rhesus Monkeys

Subject	Condition	Trials Completed	% Correct	z
Murph	Meaningful	2,766	50	.99
	Nonmeaningful	779	48	.61
Lou	Meaningful	1,826	51	.94
	Nonmeaningful	3,599	48	-2.15*
Willie	Meaningful	3,242	50	.21
	Nonmeaningful	1,100	48	-.78
Gale	Meaningful	1,306	49	-.44
	Nonmeaningful	1,174	47	-1.40
Hank	Meaningful	2,400	50	.08
	Nonmeaningful	1,178	52	1.75

* $p < 0.05$

Note. Number of trials for each individual and condition varies as a result of how many trials each monkey was able to complete in a given testing session. These data represent approximately ten sessions over the span of two weeks.
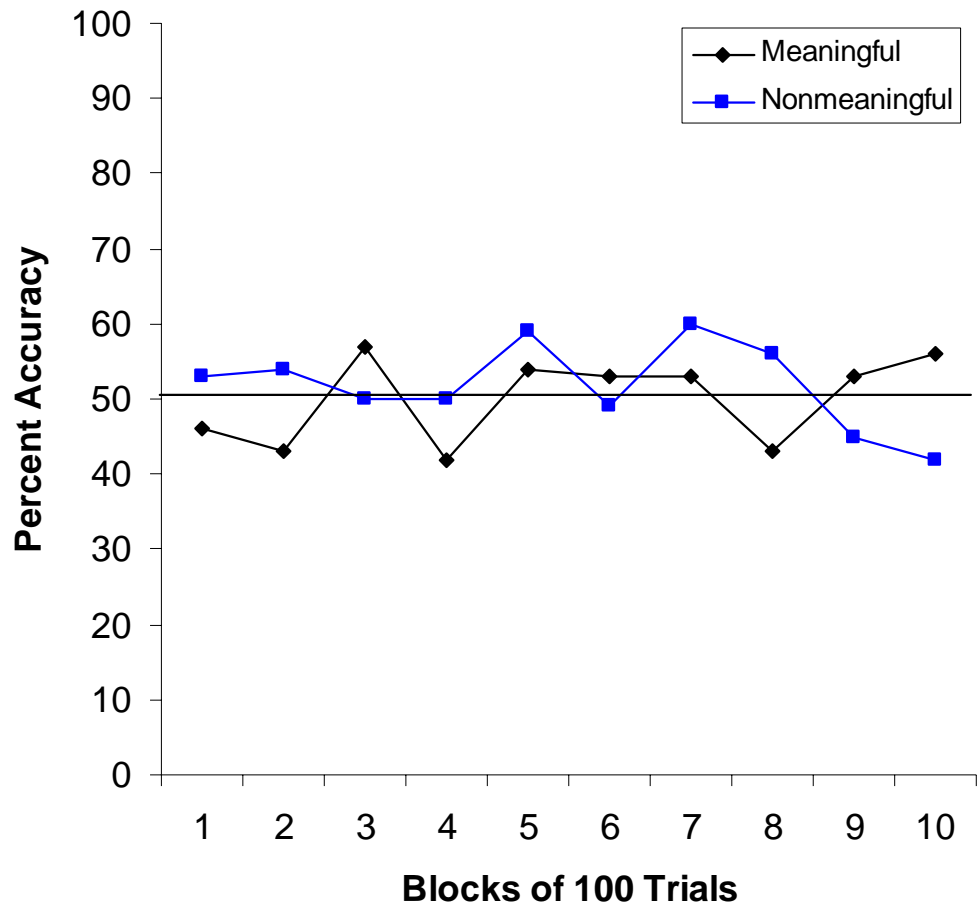
Figure 13. Percent Correct by 100-Trial Block for Hank

Table 3.

Analysis by 100-Trial Blocks for Hank

	Meaningful		Nonmeaningful	
	% correct	z	% correct	z
1-100	46	-0.8	53	0.6
101-200	43	-1.4	54	0.8
201-300	57	1.4	50	0.0
301-400	42	-1.6	50	0.0
401-500	54	0.8	59	1.8
501-600	53	0.6	49	-0.2
601-700	53	0.6	60	2.0*
701-800	43	-1.4	56	1.2
801-900	53	0.6	45	-1.0
901-1000	56	1.2	42	-1.6

* $p < 0.05$

All but one monkey completed *same* and *different* trials with equivalent levels of accuracy in every condition; their treatment of *same* and *different* trials was symmetric (Murph-numerals χ^2 (1, 2766) = .92, *p* > .05; Murph-letters χ^2 (1, 779) = .24, *p* > .05; Lou-numerals χ^2 (1, 1826) = .07, *p* > .05; Lou-letters χ^2 (1, 3599) = .03, *p* > .05; Willie-numerals χ^2 (1, 3242) = .87, *p* > .05; Willie-letters χ^2 (1, 1100) = .05, *p* > .05; Gale-numerals χ^2 (1, 1306) = .23, *p* > .05; Gale-letters χ^2 (1, 1174) = 2.43, *p* > .05; Hank-numerals χ^2 (1, 2399) = 2.12, *p* > .05). In the nonmeaningful condition, Hank completed same trials (56.12%) with significantly greater accuracy than different trials (48.70%) (χ^2 (1, 1178) = 6.60, *p* > .05).

As with the chimpanzees, positional biases also tended to shift between conditions for each animal. For instance, if the animal developed a right bias in the nonmeaningful condition, in the subsequent meaningful condition, a left-side bias emerged. After prolonged failure, all side biases were evident within the first 500 trials. Table 4 summarizes position biases observed for rhesus monkeys.

Discussion

Humans

For human participants, the known, discrete meaning of stimuli appeared to facilitate the acquisition of an analogical rule. Although performance was generally high across conditions, participants in the meaningful condition performed at significantly higher rates of accuracy than those in the nonmeaningful condition. For those participants in the meaningful condition, the analogical rule was also learned in fewer trials, albeit this difference was not statistically significant.

Table 4.

Position Biases During First 500 Trials for Rhesus Monkeys

Subject	Condition	Side Biased	% Biased	z
Murph	Meaningful	Left	74	24.79*
	Nonmeaningful	Right	88	21.67*
Lou	Meaningful	Left	94	40.99*
	Nonmeaningful	Right	97	56.89*
Willie	Meaningful	Left	92	47.91*
	Nonmeaningful	Right	92	27.68*
Gale	Meaningful	Left	76	18.54*
	Nonmeaningful	Right	81	21.30*
Hank	Meaningful	Left	87	36.66*
	Nonmeaningful	Right	75	16.96*

* $p < 0.01$

Response times, although significantly shorter after criterion was met, did not illustrate any effect meaningful stimuli. Differences in response time pre- and post-acquisition may reflect a practice effect as a result of increased joystick familiarity over time. Whereas this may be a confound of the measure, a more parsimonious explanation is that participants spent more time examining stimuli and attempting to determine the rule on trials before criterion was achieved.

The representational value of each stimulus enabled the relational concepts of sameness and difference to be more salient to the participants in the meaningful condition than those who completed the task with (nonmeaningful) stimuli that had no inherent representational value. The discrete meaningful value of a stimulus not only enhances its own uniqueness, but may also remove associations it may have to the stimulus with which it is paired. For instance, the word "apple" always has the same representational value, but did not always appear paired with the word "carrot" that has its own distinct meaning. This dual role for meaning may allow for relational information about a pair of stimuli to emerge as the salient overall stimulus quality in fewer trials than if specific perceptual dimensions of a stimulus must be closely attended to on each trial.

It is interesting that some people failed to learn the analogical rule under any condition. On their de-briefing forms, these participants noted that they simply "never figured it out" or "tried to match similar things, but that didn't work." Therefore, relational similarities and differences never became salient as part of a rule-learning strategy. Like children, (Gentner, 1988) they may have been distracted by the surface similarities between the components of one trial. Whereas failure to learn the analogical rule may have been due to differences in motivation to participate, it is likely that if

structural similarities (i.e., the matching rule) were disclosed to the participants at the beginning of the experiment, success would be imminent.

Chimpanzees

One chimpanzee showed above chance performance on the ago relational s/d task, but only in the meaningful-lexigram condition. A second chimpanzee also performed at levels above chance, but in the opposite condition (nonmeaningful-lexigram). Individual differences in performance by the chimpanzees may have arisen for a variety of reasons: the individual does not have the capacity for analogical reasoning, the representational value of the lexigram (or numeral) is not strong enough to differentiate it and allow for the emergence of a relational rule for the individual, or the representational value of the meaningful stimuli was so strong that the animal did not have the ability to overcome its inherent value and recognize relational information within a pair of these meaningful stimuli. In addition, asymmetric effects on *same* versus *different* trials may reflect an inflexibly acquired rule. Asymmetric effects are an indication that half of the analogical rule is understood, but that the opposite of that rule may not also hold true.

For Lana and Mercury, whose performance did not differ significantly from chance in any condition, meaning most likely did not function to make the stimuli discrete entities, but rather the pairing of stimuli functioned only as a set without independent constituent parts. Because these animals were not able to extract relational information from the sample pair, completing the analogical problem correctly was impossible.

Panzee, who successfully completed only trials in the meaningful condition, and not nonmeaningful or number, seems dependent on the meaningfulness of stimuli in

order to match the choice relation to the sample. One should note that whereas this difference between meaningful (lexigrams) and nonmeaningful conditions is significant, an accuracy of 66% for the meaningful (lexigram) condition is less convincing than a criterion of 80% or better that is typically regarded as evidence for task acquisition (e.g., Blaisdell & Cook, 2005; Cook, Cavoto, & Cavoto, 1995; Wasserman et al., 2002; Young & Wasserman, 2002). Rather than facilitating the acquisition of the analogical rule, one chimpanzee failed to acquire the analogical rule at all unless the stimuli have discreet meaning. Even after completing the meaningful sessions first, Panzee failed to perform above levels of chance during nonmeaningful and number sessions. This would indicate that what she learned during the meaningful sessions was not a general overarching analogical rule as we might conceive it. If that were the case, she should have transferred her knowledge from the first condition (meaningful) to the next conditions presented. Rather, it is possible that she viewed the nonmeaningful and numeral conditions as completely novel tasks that she had to learn. Alternatively, the analogical rule that she learned may be specific to those stimuli with external representation and thus the rule could not be applied in instances in which meaningful symbols are not present. However, by this logic, she should have been able correctly to complete trials with Arabic numerals present. That she did not perform above levels of chance with these stimuli either may indicate that Arabic numerals may not represented in as concrete a way that objects such as foods and places are (i.e., representation as estimates; see Beran, 2004) . Perhaps her analogical rule applied in a rather limited way only to symbols for specific, exactly-defined objects.

Sherman responded in a manner opposite to Panzee. He completed trials only in the nonmeaningful condition above levels of chance. Again, whereas this difference between meaningful (lexigrams) and nonmeaningful conditions is significant, an accuracy of 65% in the nonmeaningful condition does not represent a very strong effect. Rather than assisting his acquisition of an analogical rule, meaning may act as a confusing factor for Sherman. That he could correctly complete only nonmeaningful trials may indicate that his analogical rule was more perceptually than symbolically based. Perhaps he attended more specifically to the perceptual qualities of a stimulus in order to determine its relatedness to another within the pair. In the meaningful and *number* conditions, it may be the case that when a meaningful stimulus was present, Sherman did not look past the specific meaning associated with it in order to search for relational information. Rather, Sherman may search only for stimuli that match in their meaning and not in their relational value. For both Panzee and Sherman, poor performance on subsequent conditions may be indicative of the inflexibility of their application of the analogical rule.

Asymmetric performance by Sherman on *different* trials in comparison to *same* trials in the nonmeaningful condition indicates that the analogical rule acquired applied only to differently related pairs of items. Rather than using the rule to match relations flexibly across trial types, asymmetric performance indicates that the strategy used was not fully developed so that a consistent rule could be applied to *same* relations.

Both pigeons and people are predisposed to notice differences as well as similarities (Fagot et al., 2001; Young & Wasserman, 2002). Like pigeons and people, chimpanzees exhibit some differences in performance on *same* versus *different* trials. If

we suppose that chimpanzees are predisposed to noticing only differences, it is likely that (at least for Sherman) an analogical rule was realized only through experience matching differently related pairs with other differently related pairs. In subsequent trials, when similarly related pairs were presenting, the limited analogical rule that he learned no longer applied.

Monkeys

Failure to match relational pairs correctly in this task could be the result of one or more of a variety of reasons: the monkeys cannot extract the necessary relational information from a pair of objects, relational knowledge is not encoded such that it is accessible for application to novel behaviors, or perceptual properties of stimuli can not be ignored in a matching paradigm.

From their performance on previous tasks (e.g., Flemming et al., 2005), we know that monkeys can extract relational information from a pair of objects. In a two-choice discrimination paradigm, monkeys chose either a pair of identical or nonidentical objects in the presence of a discriminative cue. It may be the case that the monkey's ability to extract relational information is reliant on a discriminative cue; the discriminative cue prompts the search for relational information present in the given sample. In the current task, no discriminative cue was offered, perhaps not enabling the search for a relation between the items in the pair.

Whereas monkeys possess the ability to extract relational information from a pair of items in certain testing instances, they fail to express this type of knowledge in a relational match-to-sample (MTS) paradigm. Their history with the MTS paradigm may be so closely tied the task to the search for perceptual similarities amongst presented

items that requesting a different rule (the search for relational similarities) is unreasonable to expect. When pairs of items are presented, and no choice pair includes similar perceptual qualities, monkeys may be inflexible in their quest for a new rule to impose on the seemingly familiar task.

Relational matching may not be beyond the capability of rhesus monkeys or some chimpanzees, but it would seem that relational information is certainly not as salient as it seems to be for humans. Presenting the task in a slightly modified paradigm could make more salient the relational information to be utilized in the task. One plausible variation could present each stimulus individually and successively, forcing attention first to the uniqueness of each item composing the pair, followed by combined attention to the pair itself. In addition, removal of the sample pair before choice pair presentation can reduce visual complexity on the screen. If the sample pair is removed from the screen, the search for a perceptually identical match can be diminished. In this new task variation, attention to the pairs as a grouping of two objects rather than a conglomerate, and a diversion from perceptual matching are achieved.

General Discussion

Robust differences in performance emerged across species in this analogical reasoning task. Not only did human participants outperform chimpanzees and monkeys, but the role of meaning when completing an analogy was also dissimilar across species. According to these data, stimuli with representational value can facilitate, hinder, or have no effect on the completion of an analogy of *same* and *different* objects.

Taken together, the results of this comparative study both lend support and opposition to the previous suggestion that language-like abilities and symbolic training are integral to the capacity of analogical reasoning. As was discussed above, Premack (1983) concluded that language-like training is necessary based on his experiments with Sarah, a chimpanzee. Similarly, Thompson et al. (1997) posited that labeling of relational information is a necessary component of analogical thinking. Whereas the results of the current study do not allow for the conclusion that human and chimpanzee participants were labeling the related pairs of stimuli, individual differences which arose between the chimpanzees urge the consideration of an alternate hypothesis. With extensive lexical vocabularies, three of the four chimpanzees that participated in this study have the capacity to label items, and by Thompson and colleagues' (1997) logic should possess the ability to conceptualize relations-between-relations. One chimpanzee was not able to do so in any of three conditions (meaningful, nonmeaningful, and number). One chimpanzee was able to complete the task only when individual stimuli afforded external representation (meaningful); still another chimpanzee successfully completed the task only when stimuli were unfamiliar and nonmeaningful. These differences in performance suggest one of three alternate hypotheses related to labeling: labeling of relations is dependent upon different stimulus qualities for different individuals, labels must exist for specific *same* and *different* terms, or labeling plays no role in analogical reasoning.

Whereas one chimpanzee has a lexical vocabulary and can certainly label external objects with corresponding lexigrams, above chance performance on the analogical reasoning task would be expected, but was not observed. This poor performance may mean that labeling, if in fact occurring during the task, plays no role in the acquisition of

an analogical rule. Another possibility, though, is that labeling of relational information is not made more salient than the overt object labels (of individual stimuli) themselves. That is, meaning of individual stimuli confounds the labeling of the related pair for this individual. This could also be the case for the chimpanzee that correctly completed trials in the nonmeaningful condition, but not in meaningful or number. Perhaps after correctly completing trials in the nonmeaningful condition (either with or without succinct labels or re-coding terms of relatedness), this individual had used labels that were specific to physical similarities between stimuli to recode the pair. In the meaningful condition, identity pairs are physical matches in addition to their matched meaning. Perhaps when representational value for specific external objects is present, that meaning becomes the default dimension of salience, rather than allowing for the search of abstract labels that are not overtly present, but must be created by attending only to the relatedness of stimuli.

The numeral condition carried out with the chimpanzees deserves special consideration. Numerical studies with the chimpanzees (e.g. Beran, 2004) suggest that Arabic numerals, like lexigrams, carry with them some type of representational value. How these numerals are represented is yet undetermined. Rather than concrete quantity information, it is likely that these Arabic numerals represent a vague estimate of a number of items (Beran, 2004). For this reason, although they are meaningful, they should not be placed on a level of representational capacity equal to lexigrams. Lexigrams are symbols that are used on a daily basis to reference specific people, places, objects, and foods. Numerals, however, are used only in a small percentage of cognitive tasks, and are not spontaneously used by the chimpanzees to reference, for example,

"how many" of an item they would like to obtain. Therefore, their concept of number, aside from the little experimenters know about it, is surely less broad than the knowledge the chimpanzees have of the concrete objects they commonly reference with lexigrams.

What Analogies Mean for the S/D Concept

The relational matching paradigm utilized throughout this study tests analogical reasoning skills by forcing explicit expression of conceptual knowledge for same and differently related pairs. Thus, analogy completion assumes (and relies on) conceptual knowledge of same and different. Certainly, human participants in this study have a very broad conception of same and different, as humans regularly classify objects into groups based on similarities and differences. Perhaps our propensity to do so is driven by the salience of relations. Whereas it is clear that other animals have the capacity to perceive the relations of same and different (Bovet & Vauclair, 2001; Cook et al., 1995; Flemming et al. 2005; Katz et al., 2002; Wasserman et al., 2002), the acquisition of these concepts for birds and monkeys does not emerge for sometimes hundreds of trials. Human participants as young as 3 years old have provided evidence that the identity/nonidentity concept emerges in significantly fewer trials (Gentner & Markman, 1997).

Because the concepts of *same* and *different* are not as salient to nonhuman primates as they are to humans, their use when searching for an already abstract matching rule (in RMTS) is not as readily available a strategy. Pigeons and monkeys which have been shown in some studies to rely on arrays of multiple items in order to glean relational information (e.g., Wasserman et al., 2002) may rely on more ecological valid strategies (i.e., colors and shapes) when presented with a matching to sample task. Unlike simple discriminations that rely only on implicit types of conceptual knowledge, the analogical

paradigm carried throughout this study requires the explicit use of the same/different strategy. In sum, because the same/different concept is less salient and possibly more narrowly conceived for nonhuman primates, it stands to reason that the application of the concept would be more difficult for such an animal.

Species Differences: Why Analogies Are Difficult Even for Us

Application of analogical reasoning sometimes *can* occur spontaneously for humans, rather than explicitly being told to use an analogy in order to solve a problem (Dunbar, 2001). In most scenarios, like standardized tests, we are instructed explicitly to think of the problem in terms of relationships and that analogical reasoning should be applied. This type of instruction is impossible to deliver to a nonhuman primate. That the chimpanzees in this study complete the problem at all (without explicit instruction) is impressive.

When the similarities between old and new problems are surface, (e.g., using the same elements) analogical reasoning is more quickly applied to the new problem. Rather, if the similarities between previous experiences and novel problems are only structural (e.g. content) in nature, application of analogical reasoning skills is far less obvious and therefore not as salient a strategy.

In experiments with children and adult humans, researchers agree that surface similarities are the key to whether participants will think of using an analogy to solve a problem when not explicitly told to do so (Catrambone, 2002; Gentner, Rattermann, & Forbes, 1993). In addition, human participants are particularly distracted by surface similarities in analogous problems, even when they are unimportant (Ross, 1987).

In the current study, only structural similarities exist across trials. Surface similarities in the current paradigm would only exist if individual stimuli were consistently arranged across trials. The repeated presentation of individual stimuli (albeit in various locations and combinations of pairings) could be interpreted as surface similarities which act as confounds not related to the application of analogical knowledge. If stimulus A is present in the sample pair, not only may the subject attempt to search for stimulus A amongst the choice pairs, but they may also retain knowledge of the role of stimulus A for future trials. Thus, if stimulus A is encountered again, regardless of current location or pairing, rules previously associated with it may be incorrectly applied to the current scenario.

Children as old as 11 years old often have difficulties in their analogical reasoning abilities. Rather than responding on the basis of relational similarities, thematic- and object-similarity choices are often are more salient to children (Rattermann & Gentner, 1998). Later in development, most children successfully solve analogical reasoning noting relational similarities and differences. For this reason, there is likely a shift whereby children interpret analogy only in terms of object similarity and then in terms of relational similarity (Gentner, 1988; Ratterman & Gentner, 1998).

In sum, the present study uncovers both common threads and disparities in the analogical reasoning skills of members of the primate lineage. Whereas meaningful stimuli act as facilitators for humans in making more salient the relational information presented and consequently the nature of an analogical rule, meaning can take on various roles for other primate species. Representational value can take on the role of both enabling and inhibiting analogy completion for chimpanzees, and seems to have no

facillatory role at all for rhesus monkeys. These different roles may reflect the different ways that nonhuman primates represent their worlds. Despite the ambiguity of the exact mechanism by which analogies are realized, the present experiment provides the first parallel comparison of the acquisition of an analogical rule across three species of primates.

References

Beran, M. J. (2004). Chimpanzees (*Pan troglodytes*) respond to nonvisible sets after one-by-one addition and removal of items. *Journal of Comparative Psychology, 118,* 25-36.

Beran, M. J., Pate, J. L., Richardson, W. K., Rumbaugh, D. M. (2000). A chimpanzee's (*Pan troglodytes*) long-term retention of lexigrams. *Animal Learning and Behavior, 28,* 201- 207.

Beran, M. J. & Washburn, D. A. (2002). Chimpanzee responding during matching to sample: Control by exclusion. *Journal of the Experimental Analysis of Behavior, 78,* 497-508.

Bovet, D., & Vauclair, J. (2001). Judgment of conceptual identity in monkeys. *Psychonomic Bulletin & Review, 8,* 470-475.

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

Catrambone, R. (2002). The effects of surface and structural feature matches on the access of story analogs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 318-334.

Cook, R. G., Cavoto, K. K., & Cavoto, B. R. (1995). Same-different texture discrimination andconcept learing by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 21,* 253-260.

Cook, R. G., Katz, J. S., & Cavoto, B. R. (1997). Pigeon same-different concept learning with multiple stimulus classes. *Journal of Experimental Psychology: Animal Behavior Processes, 23,* 417-433.

Cook, R. G., Wright, A. A., & Kendrick, D. F. (1990). Visual categorization in pigeons. In M. L. Commons, R. Herrnstein, S. M. Kosslyn, & D. B. Mumford (Eds.), *Quantitative analyses of behavior: Behavioral approaches to pattern recognition and concept formation* (pp. 187-214). Hillsdale, NJ: Erlbaum.

D'Amato, M. R. & Van Sant, P. (1988). The person concept in monkeys (*Cebus paella*). *Journal of Experimental Psychology: Animal Behavior Processes, 14,* 43-55.

Dunbar, K. (2001). The analogical paradox: Why analogy is so easy in naturalistic settings, yet so difficult in the psychological laboratory. In D. Gentner, K. J. Holyoak, & B. Kokinov (Eds.), *Analogy: Perspectives from cognitive science* (pp. 313-334). Cambridge, MA: MIT Press.

Duyck, W., Desmet, T., Verbeke, L., & Brysbaert, M. (2004). WordGen: A Tool for Word Selection and Non-Word Generation in Dutch, German, English, and French. *Behavior Research Methods, Instruments & Computers, 36,* 488-499.

Engelmann, S. (1969). *Conceptual Learning.* San Rafael, California: Dimensions Publishing Co.

Fagot, J., Wasserman, E. A., & Young, M. E. (2001). Discriminating the relation between relations: The role of entropy in abstract conceptualization by baboons (*Papio papio*) and humans (*Homo sapiens*). *Journal of Experimental Psychology: Animal Behavior Processes, 27,* 316-328.

Flemming, T. M., Beran, M. J., & Washburn, D. A. (2005, March). Same vs. different or uniformity vs. chaos?: Perception of abstract relations by Rhesus monkeys (*Macaca mulatta*). Poster session presented at the annual meeting of the Comparative Cognition Society, Melbourne, FL.

Fortes, A. F., Merchant, H., & Georgopoulos, A. P. (2004). Comparative and categorical

spatial judgments in the monkey: "high" and "low." *Animal Cognition, 7,* 101-

108.

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

Development, 59,* 47-59.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity.

American Psychologist, 52, 45-56.

Gentner, D., Rattermann, M. J., & Forbes, K. D. (1993). The roles of similarity in

transfer: Separating retrievability from inferential soundness. *Cognitive

Psychology, 25,* 431-467.

Gillian, D. J., Premack, D., & Woodruff, G. (1981). Reasoning in the chimpanzee: I.

Analogical reasoning. *Journal of Experimental Psychology: Animal Behavior

Processes, 7,* 1-17.

Goldstone, R. L. & Barsalou, L. W. (1998). Reuniting perception and conception.

Cognition, 65, 231-262.

Halford, W. K., & Graeme, S. (1992). Analogical reasoning and conceptual complexity

in cognitive development. *Human Development, 35,* 193-217.

Harris, E. H. & Washburn, D. A. (2005). Macaques' (Macaca mulatta) use of numerical

cues in maze trials. *Animal Cognition, 8,* 190-199.

Herman, L. M., Hovancik, J. R., Gory, J. D., & Bradshaw, G. L. (1989). Generalization

of visual matching by a bottlenosed dolphin (*Tursiops truncates*): Evidence for

invariance of cognitive performance with visual and auditory materials. *Journal

of Experimental Psychology: Animal Behavior Processes, 15,* 124-136.

Herrnstein, R. J. (1990). Levels of stimulus control: A functional approach. *Cognition, 37,* 133-166.

Holyoak, K. J., & Thagard, P. (1997). The analogical mind. *American Psychologist, 52,* 35-44.

Katz, J. S., Wright, A. A., & Bachevalier, J. (2002). Mechanisms of *same/different* abstract-concept learning by rhesus monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes, 28,* 358-368.

Kinnaman, A. J. (1902). The mental life of two macacus rhesus monkeys in captivity. *American Journal of Psychology, 13,* 98-148.

Kirkpatrick-Steger, K., & Wasserman, E. A. (2000). The pigeon's discrimination of shape and location information. *Visual Cognition, 7,* 417-436.

Pearce, J. M. (1994). Discrimination and categorization. In E. C. Carterette & M. P. Froedman (Series Ed.) & N. J. Mackintosh (Vol. Ed.), *Handbook of Perception and Cognition: Animal Learning and Cognition* (pp. 109-134). New York: Academic Press.

Pepperberg, I. M. (1987). Acquisition of the same/different concept by an African Grey parrot (*Psittacus erithacus*): Learning with respect to categories of color, shape, and material. *Animal Learning & Behavior, 15,* 423-432.

Premack, D. (1976). *Intelligence in ape and man.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Premack, D. (1983). Animal cognition. *Annual Review of Psychology, 34,* 351-362.

Premack, D. (1986). Minds with and without language. In L. Weiskrantz (Ed.), *Thought without tlanguage* (pp. 46-65). New York: Oxford University Press.

Premack, D., & Premack, A. (2003). *Original intelligence: unlocking the mystery of who we are.* New York: McGraw-Hill.

Rattermann, M. J. & Gentner, D. (1998). More evidence for a relational shift in the development of analogy: Children's performance on a causal mapping task. *Cognitive Development, 13,* 453-478.

Roberts, W. A., Mazmanian, D. S. (1988). Concept learning at different levels of abstraction by pigeons, monkeys, and people. *Journal of Experimental Psychology: Animal Behavior Processes, 14,* 247-260.

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

Rumbaugh, D. M. (2002). Emergents and rational behaviorism. *Eye on Psi Chi, 6,* 8-14.

Rumbaugh, D. M. & Washburn, D. A. (2003). *Intelligence of apes and other rational beings.* New Haven, CT: Yale University Press.

Rumbaugh, D. M., & Washburn, D. A. (1991). Ordinal judgments of numerical symbols by macaques (*Macaca mulatta*). *Psychological Science, 2,* 190-193.

Santiago, H.C., & Wright, A. A. (1984). Pigeon memory: Same/different concept learning, serial probe recognition acquisition and probe delay effects on the serial-position function. *Journal of Experimental Psychology, Animal Behavior Processes, 10,* 498-512.

Savage-Rumbaugh, E. S., Rumbaugh, D. M., Smith, S. T., & Lawson, J. (1980). Reference: The linguistic essential. *Science, 210,* 922-925.

Savage-Rumbaugh, E. S., Fields, W. M., Taglialatela, J. P. (2001). Language, speech, tools, and writing. *Journal of Consciousness Studies, 8,* 273-292.

Shannon, C. E., & Weaver, W. (1949). The mathematical theory of communication. Champaign, IL: University of Illinois Press.

Shields, W. E., Smith, J. D., & Wasburn, D. A. (1997). Uncertain responses by humans and rhesus monkeys (*Macaca mulatta*) in a psychophysical same-different task. *Journal of Experimental Psychology: General, 126,* 147-164.

Shyan, M. R., & Wright, A. A. (1993). The effects of language on information processing and abstract concept learning in dolphins, monkeys, and humans. In H. L. Roitblat & L. M. Herman (Eds.), *Language and communication: Comparative perspectives* (pp. 385-402). Hillsdale, NJ: Erlbaum.

Steirn, J. N., & Thomas, R. K. (1990). Comparative assessments of intelligence: Performances of *Homo sapiens sapiens* on hierarchies of oddity and sameness-difference tasks. *Journal of Comparative Psychology, 104,* 326-333.

Sternberg, R. J. (1977). Component processes in analogical reasoning. *Psychological Review, 84,* 353-378.

Thomas, R. K. (1980). Evolution of intelligence: An approach to its assessment. *Brain, Behavior, and Evolution, 17,* 454-472.

Thompson, R. K. R. (1995). Natural and relational concepts in animals. In H. L. Roitblat & J. Meyer. *Comparative Approaches to Cognitive Science* (pp. 175-224). Cambridge, Massachusetts: MIT Press.

Thompson, R. K. R., & Oden, D. L. (1996). A profound disparity revisited: Perception and judgment of abstract identity relations by chimpanzees, human infants, and monkeys. *Behavioral Processes, 35,* 149-161.

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

Thompson, R. K. R., Oden, D. L., & Boysen, S. T. (1997). Language-naïve chimpanzees (*Pan troglodytes*) judge relations between relations in a conceptual matching-to-sample task. *Journal of Experimental Psychology: Animal Behavior Processes, 23,* 31-43.

Washburn, D. A. (1994). Stroop-like effects for monkeys and humans: Processing speed or strength of association? *Psychological Science, 5*, 375-379.

Washburn, D. A., Astur, R. S. (1998). Nonverbal working memory of humans and monkeys: Rehearsal in the sketchpad? *Memory & Cognition, 26,* 277-286.

Washburn, D. A., Rumbaugh, D. M., & Richardson, W. K. (1992). The language research center's computerized test system for environonmeaningfulental enrichment and psychological assessment. *Contemporary Topics, 31,* 11-15.

Washburn, D. A., Thompson, R. K. R., & Oden, D. L. (1997). Monkeys trained with same/different symbols do not match relations. Paper presented at the 38th Annual Meeting of the Psychonomic Society. Philadelphia, PA.

Wasserman, E. A., Frank, A. J., & Young, M. E. (2002). Stimulus control by same-versus-different relations among multiple visual stimuli. *Journal of Experimental Psychology: Animal Behavior Processes, 28,* 347-357.

Wasserman, E. A., Young, M. E., & Fagot, J. (2001). Effects of number of items on the baboon's discrimination of same from different visual displays. *Animal Cognition, 4,* 163-170.

Wasserman, E. A., Young, M. E., & Peissig, J. J. (2002). Brief presentations are sufficient for pigeons to discriminate arrays of same and different stimuli. *Journal of the Experimental Analysis of Behavior, 78,* 365-373.

Wright, A. A. (1997). Concept learning and learning strategies. *Psychological Science, 8,* 119-123.

Wright, A. A., Rivera, J. J., & Katz, J. S. (2003). Abstract-concept learning and list-memory processing by capuchin and rhesus monkeys. *Journal of Experimental Psychology: Animal Behavior Processes, 29,* 184-198.

Wright, A. A., & Santiago, H. C. (1984). Monkey memory: Same/different concept learning, serial probe acquisition, and probe delay effects. *Journal of Experimental Psychology: Animal Behavior Processes, 10,* 513-529.

Young, M. E., & Wasserman, E. A. (1997). Entropy detection by pigeons: Response to mixed visual displays after same-different discrimination training. *Journal of Experimental Psychology: Animal Behavior Processes, 23,* 157-170.

Young, M. E., & Wasserman, E. A. (2002). Detecting variety: What's so special about uniformity? *Journal of Experimental Psychology: General, 131,* 131-143.

Zentall, T. R., Galizio, M., & Critchfield, T. S. (2002). Categorization, concept learning, and behavior analysis: An introduction. *Journal of the Experimental Analysis of Behavior, 78,* 237-248.