What Meaning Means for Same and Different: Analogical Reasoning in Humans (*Homo sapiens*), Chimpanzees (*Pan troglodytes*), and Rhesus Monkeys (*Macaca mulatta*)

Timothy M. Flemming,
Department of Psychology and the Language Research Center, Georgia State University

Michael J. Beran,
Language Research Center, Georgia State University

Roger K. R. Thompson,
Department of Psychology, Franklin&Marshall College

Heather M. Kleider, and
Department of Psychology, Georgia State University

David A. Washburn
Department of Psychology and the Language Research Center, Georgia State University

Abstract

Thus far, language- and token-trained apes (e.g., D. Premack, 1976; R. K. R. Thompson, D. L. Oden, & S. T. Boysen, 1997) have provided the best evidence that nonhuman animals can solve, complete, and construct analogies, thus implicating symbolic representation as the mechanism enabling the phenomenon. In this study, the authors examined the role of stimulus meaning in the analogical reasoning abilities of three different primate species. Humans (*Homo sapiens*), chimpanzees (*Pan troglodytes*), and rhesus monkeys (*Macaca mulatta*) completed the same relational matching-to-sample (RMTS) tasks with both meaningful and nonmeaningful stimuli. This discrimination of relations-between-relations serves as the basis for analogical reasoning. Meaningfulness facilitated the acquisition of analogical matching for human participants, whereas individual differences among the chimpanzees suggest that meaning can either enable or hinder their ability to complete analogies. Rhesus monkeys did not succeed in the RMTS task regardless of stimulus meaning, suggesting that their ability to reason analogically, if present at all, may be dependent on a dimension other than the representational value of stimuli.

Keywords

meaning; representation; same–different; analogical reasoning; primate

Analogical reasoning, Halford and Graeme (1992) argued, is the mechanism that allows for conceptual thinking, including logical inference. Knowledge about analogies forces explicit

expression of conceptual knowledge, unlike simple discriminations that rely on implicit types of conceptual knowledge (Premack, 1976). In an analogy, a relationship must be established between the first two elements in the series (i.e., the base relation). Then, and only then, can one continue to the second set of elements (i.e., the target relation) and seek the same relation instantiated in the base (Gentner, Ratterman, & Forbes, 1993; Thompson & Oden, 2000). If one matches two abstract relations, one has in essence recognized the analogical equivalence of relations.

Evidence of this ability is very rare in nonhuman animals, although the chimpanzee Sarah provided perhaps the best evidence of analogical reasoning (Gillian, Premack, & Woodruff, 1981; Premack, 1983; Oden, Thompson, & Premack, 2001). 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 (color, shape, or size) were placed to the left of a center chip that signified the concept *same*. This plastic chip with specialized symbolic meaning seems to be the impetus by which Sarah was capable of judging the relations between those relations. 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 did Sarah complete the task with flat geometric shapes, but she was also successful when the items presented were everyday three-dimensional objects (Gillian et al., 1981).

Whereas it is often difficult for humans to describe identity–nonidentity relations, for which nonhuman primates possess some knowledge, without using the words *same* and *different*, these abstract concepts have no linguistic prerequisites (Premack, 1976). These words can be applied to objects that themselves do not have names. Moreover, in relation-level problem solving, the names of the objects are 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, 1986; Thompson & Oden, 1996).

Premack (1976, 1983, 1986) suggested that at least in the case of abstract relations, acquisition of conceptual knowledge depends on language. Accordingly, it should come as no surprise that although 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 that humans understand them (Premack, 1983, 1986; Thompson, 1995; Thompson & Oden, 2000). Symbolic representation of relations has further been implicated as the main mechanism responsible for the judgment of relations-between-relations—that is, second-order relations (Premack, 1976, 1983; Thompson & Oden, 2000).

As Thompson and Oden (1996, 2000) posited, to complete a relational matching-to-sample problem successfully, participants must recognize the relationship between the members of each pair as the critical aspects of the problem. How might one go about recognizing these relations? Acquisition of concrete symbolic tokens of otherwise abstract relationships may be the necessary (and perhaps sufficient) component to any relational matching-to-sample (RMTS) task (Oden et al., 2001). Thompson, Oden, and Boysen (1997) hypothesized that

the judgment of relations-between-relations is made possible by an animal's representational capacity to reencode abstract relations into iconically equivalent symbols. They presented adult chimpanzees with a RMTS task after they had learned to choose a heart-shaped token when presented with a pair of identical objects and to choose a diagonally shaped token when presented with a nonidentical pair of objects. In the RMTS task, the chimpanzees viewed paired random junk objects as identical or nonidentical sample pairs and pictorial paired stimuli on a touch screen as the target choice stimuli. The goal was to indicate the target that conveyed the same relation between the objects as in the base sample. Four of five chimpanzees spontaneously judged the conceptual equivalence of relations-between-relations in the absence of the symbolic tokens. The fifth chimpanzee differed in his learning history; unlike the other chimpanzees, he was naïve with respect to both numeric problem solving and symbolic token training. Therefore, it seems that this experience with symbols may have played a functional role in the acquisition of abstract concepts in these chimpanzees. Thompson and Oden (1996) suggested that the critical role of these symbols was to provide animals with a concrete icon for encoding a propositional representation that is otherwise abstract. Thompson et al. (1997) also suggested that conceptual–relational matching is akin to covert symbol matching. It should follow, then, that such symbolic training produces a system for universal computation (Clark & Thornton, 1997).

Can abstract relations be labeled without symbolic representation? If the answer to this question is "no," then we should not be surprised by Premack's (1983) 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, and in particular symbolic representation, then we should expect that language-naïve nonhuman species could also succeed on tasks requiring analogical reasoning, if they have the truly relevant cognitive mechanisms necessary for such reasoning.

Acquisition of, and discrimination between, abstract relations has been investigated in several species of Old-World monkeys. Fagot, Wasserman, and Young (2001), for example, demonstrated that at least two baboons had marked difficulty with an identity–nonidentity RMTS task with multiple icons. Multiple item displays were used as stimuli in a matching-to-sample task. One stimulus array (composed of 16 or fewer identical or nonidentical items) served as the sample. One choice pair was composed of identical images (but different from any of those present in the sample pair) and the other was made of numerous images that differed physically from each other. Baboons were required to choose the array that matched the relation (either same or different) of the sample array. Baboons completed the task successfully until stimulus arrays contained 4 or fewer items, providing evidence that entropy detection most likely underlaid same–different discrimination abilities (Fagot et al., 2001). Thus, the relational matching strategy in this task was more perceptually bound.

By contrast, some judgments of identity are conceptually bound. Baboons also proved proficient in the judgment of conceptual identity. Bovet and Vauclair (2001) trained baboons to discriminate between food and nonfood stimuli. In tests of conceptual matching, baboons were able to identify pairs of food or nonfood items as same or different on the basis of their conceptual relationships. For instance, if one apple and one banana were presented, baboons

responded by pulling the rope indicating *same* because they both belong to the food category. It is important to note that those judgments, in contrast to ones made by baboons in the Fagot et al. (2001) study, could not be made on the basis of physical similarities because "sameness" in that task required that baboons generalize among different items within a category. In addition, these types of discriminations are exemplary of first-order relations only, as opposed to second-order relations as outlined by Thompson and Oden (2000).

Flemming, Beran, and Washburn (2007) investigated the role that entropy measures, discriminative cues, and symbolic knowledge play for rhesus monkeys in the acquisition of the concepts of *same* and *different* in a computerized RMTS task. After repeatedly failing to discriminate relations between pairs of stimuli in a two-choice discrimination paradigm, monkeys rapidly learned to discriminate between eight-element arrays. Subsequent tests with smaller arrays, however, suggested that although important for the initial acquisition of the concept, entropy is not a variable on which monkeys are dependent. Not only did the rhesus monkeys choose a corresponding equivalent relational pair in the presence of a discriminative cue, but they also chose the cue itself in the presence of the relational pair— in essence, labeling those relations. Subsequent persistent failure in the judgment of relations-between-relations in the RMTS task (a nonperceptually based matching strategy), however, suggested that perhaps an as yet unidentified qualitatively different cognitive component exists that prevents monkeys from behaving analogically.

The purpose of the current study was to determine the role of stimulus meaning in the analogical reasoning abilities of three primate species. Such comparisons outline further our understanding of the cognitive capacities of various species and further our knowledge about concepts and mechanisms of concept learning in general. By implementing a more conceptual variable, meaning, into perceptually based judgments, one might expect conceptual strategies for the task to emerge more quickly than they might otherwise.

Rhesus monkeys, chimpanzees, and human participants completed an identity–nonidentity RMTS task composed of meaningful and nonmeaningful stimuli. It is important to discern what we mean by *meaningful*. We define meaningful stimuli as those that function as a cue to evoke representations of external objects or concepts. Meaningful stimuli in this experiment, unlike linguistic equivalents in previous research (e.g., Premack, 1976; Thompson et al., 1997) did not represent relational categories such as *same* and *different*, but rather referenced concrete objects or numerical quantities, making more overt the relational concept of which they are a part. Without using a symbolic token for the relational concept itself, we were further able to investigate the role of earlier "linguistic" prerequisites among the three primate species.

Method

Participants

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

Eighty-two undergraduates (67 women and 15 men) were recruited from Georgia State University's psychology research pool, with half assigned to each stimulus condition (meaningful and nonmeaningful stimuli) in the RMTS 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 on 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 matching-to-sample paradigm with joysticks (Beran & Washburn, 2002) but were naïve to the specific testing procedures involving the *same–different* concepts. Individuals were randomly assigned to receive one of two conditions (meaningful and nonmeaningful lexigram stimuli) first, followed by completion of the remaining condition. For 3 of the 4 chimpanzees, the meaningful stimuli were the lexigram symbols that they learned when young (Brakke & Savage-Rumbaugh, 1995, 1996; Rumbaugh, 1977; Savage-Rumbaugh, 1986) and have used for all of their lives. The fourth chimpanzee, Mercury, received no lexical training at all; therefore, both conditions were essentially non-meaningful, providing an additional control. 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 sorted lexigrams into labeled groups more accurately than they did real-world objects and photographs (Savage-Rumbaugh, Rumbaugh, Smith, & Lawson, 1980). The chimpanzees were not food or water deprived. Individuals worked at mobile testing systems at their home cages for designated 1-hr sessions each day.

Five male rhesus monkeys (10 to 20 years of age) housed at the LRC also completed the RMTS task. All monkeys began this study with previous experience on other tasks involving the *same–different* paradigm. In the Flemming et al. (2007) study, monkeys discriminated between arrays composed of identical or nonidentical clip-art images. In the presence of a discriminative color cue, monkeys successfully discriminated same from different displays of eight, six, four, and two items. In addition, monkeys completed but failed RMTS tasks similar to those in the current study when only nonmeaningful clip art images were used. The monkeys were not food or water deprived for purposes of testing and were allowed to work ad libitum throughout the day in their home cages.

Stimuli

Meaningful stimuli for human participants consisted of three- to seven-letter words that referred to concrete objects such as foods and places. Nonmeaningful stimuli were composed of three- to seven-letter strings of illegal nonwords generated by the program WordGen (Duyck, Desment, Verbeke, & Brysbaert, 2004). We used illegal nonwords as nonmeaningful stimuli because they carry with them no inherent referential value, and they cannot be re-coded into sensible phonemes in the English language. Both sets of stimuli appeared as white letters inside a black rectangle. Figure 1a and 1b provides example displays presented to human participants. A total of 50 unique words or 50 unique nonwords were presented throughout a testing session.

For chimpanzees, only lexigrams with which they had had extensive experience and for which they have been shown to retain meaning were presented during the meaningful condition. 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 stimuli for chimpanzees (see Figure 1c and 1d). One chimpanzee, Mercury, had no such experience, and therefore for him both conditions consisted of nonmeaningful stimuli. Each individual had a unique vocabulary subset. Consequently, the number of unique known lexigrams differed across individuals, and we used only concrete lexigrams that were part of each chimpanzee's larger unique vocabulary (Panzee = 65, Sherman = 28, and Lana = 19; Mercury's stimulus sets each included 30 randomly chosen lexigrams). The number of nonmeaningful stimuli was balanced for each individual.

There is sufficient evidence that monkeys can use some types of 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, allowing the animals to associate these numeric symbols with quantity information (Washburn & Rumbaugh, 1991). Thus, we used Arabic numerals as meaningful stimuli. In addition to simple quantity information that numerals may convey, Arabic numerals have also produced Stroop-like effects in rhesus monkeys at the LRC (Washburn, 1994). Because the meaning of the numerals interferes with judgments about amount, numerals mean amounts and thus can be said to have symbolic representation for the monkeys.

These numerals also control other types of responding. Harris and Washburn (2005) presented the monkeys with series of reinforced and nonreinforced maze trials in which Arabic numerals indicated the number of reinforced mazes that could be completed before a nonreinforced maze trial was presented. Monkeys developed a pattern of responding slower on nonreinforced trials than on 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). We used these previously learned Arabic numerals in paired stimuli in the meaningful condition (see Figure 1e and 1f).

In the nonmeaningful condition, stimuli consisted of 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. Monkeys were randomly assigned to receive one of two conditions (meaningful and non-meaningful stimuli) first, followed by completion of the remaining condition.

Although there are components of both meaningful and non-meaningful stimuli that are familiar in the words, lexigrams, letters, and numerals presented to the three species (i.e., individual letters, lines, and shapes), it is the combination of individual familiar stimulus features that makes them meaningful. Pilot experiments using the same stimuli over many trials still failed to show any evidence of relational matching in rhesus monkeys even though by the end of the experiment those stimuli certainly had become familiar. Thus, familiarity

is not the issue in this study; rather, it is meaningfulness. Of course, meaningful stimuli must be familiar stimuli, and so one cannot completely dissociate these two things.

Apparatus

The LRC's Computerized Test System consists of an IBM-compatible desktop personal computer (Washburn, Rumbaugh, & Richardson, 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-in. (43.2-cm) 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 1 m 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 used with the non-human primates. 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 pellets to chimpanzees.

Task

In the RMTS task, stimuli were presented in pairs with one sample pair (base) and two choice pairs (targets). At the initiation of a trial, one pair of stimuli (either two identical or nonidentical clip art images) was centered at the top of the computer screen; stimuli were approximately 5 cm × 3 cm. Participants were required to contact this sample pair with the cursor for the choice pairs to appear. Once contact with the sample pair was made, the joystick cursor was recentered 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, and the other contained two physically nonidentical items. More important, no stimulus in the choice pairs was ever physically identical to stimuli in the sample pair, and choice stimuli were randomly assigned to position on the screen. Thus, the task required participants to either match a pair of the form AA with a pair of the form BB (and not CD) or to match a pair of the form EF with a pair of the form GH (and not JJ). Similar testing paradigms have been used frequently (Flemming et al., 2007; Premack, 1976; Thompson et al., 1997). 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 sound if incorrect). Reward was provided on correct trials for nonhuman animals and a short intertrial interval (ITI) was then presented before the sample pair for the next trial appeared. For correct choices, rewards were automatically dispensed to the animals, accompanied by a 2-s ITI. When choices were incorrect, no food

100 trials. These 6 participants were removed from analysis for trials-to-criterion and response time. Participants in the meaningful condition met the 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, $t(74) = 1.65$, $p = .10$, $d = 4.73$.

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 or nonmeaningful—after at least 130 trials in each condition.

Panzee completed trials with meaningful stimuli before receiving the nonmeaningful condition. 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 condition (55% of 227 trials, z = 1.53, $p > .05$). Figure 2 presents Panzee's percentage of correct responses as a function of trial block for both conditions.

Panzee's accuracy on *same* trials and *different* trials was statistically equivalent in the meaningful condition, $\chi^2(1, N = 202) = 2.06$, $p > .05$. In the nonmeaningful condition (which yielded overall chance performance), her accuracy was significantly higher on *different* trials (65%) than on *same* trials (49%), $\chi^2(1, N = 186) = 4.61$, $p < .05$.

Sherman completed trials with nonmeaningful stimuli before receiving the meaningful condition. He performed at levels above chance in the nonmeaningful condition, completing 65% of 263 trials correctly (z = 4.99, $p < .01$). However, he failed to reach significance in the meaningful condition (54% of 186 trials, $z = 1.17$, $p > .05$). Figure 3 presents Sherman's percentage of correct responses as a function of trial block for both conditions.

Sherman's accuracy on *same* trials and *different* trials was statistically equivalent for the meaningful condition, $\chi^2(1, N = 227) = 0.48$, $p > .05$. In the nonmeaningful condition, his accuracy was significantly higher on *different* trials (80.6%) than on *same* trials (50%), $\chi^2(1, N = 263) = 27.18$, $p > .05$.

Rhesus Monkeys

The monkeys completed an average of 1,937 trials in each condition, but no monkey achieved levels of performance significantly above chance in either condition, regardless of which was presented first. Performance summaries for the rhesus monkeys are shown in Table 2, with illustrative accuracy levels displayed in Figure 4 for Hank (although any monkey could be shown with nearly identical results). The trends of the learning curves presented in Figure 4 with Hank 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 was not the case, as no monkey ever reached a level of performance significantly different from chance. To illustrate that there was no evidence that performance improved across trials, blocks of 100 trials were analyzed independently for Hank for the first 1,000 trials in both conditions. Percentage correct was not significantly different from chance in any trial block except for Trials 601–700 (60% correct; $z = 2.0$, $p < .05$) in the nonmeaningful condition. Subsequent trial blocks rebounded to near-chance performance.

All but 1 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, $\chi^2(1, N = 2,766) = 0.92$, $p > .05$, and letters, $\chi^2(1, N = 779) = 0.24$, $p > .05$; Lou—numerals, $\chi^2(1, N = 1,826) = 0.07$, $p > .05$, and letters, $\chi^2(1, N = 3,599) = 0.03$, $p > .05$; Willie—numerals, $x^2(1, N = 3,242) = 0.87$, $p > .05$, and letters, $\chi^2(1, N = 1,100) = 0.05$, $p > .05$; Gale—numerals, $\chi^2(1, N = 1,306) = 0.23$, $p > .05$, and letters, $\chi^2(1, N = 1,174) = 2.43$, $p > .05$; and Hank—numerals, $\chi^2(1, N = 2,399) = 2.12$, $p > .05$. In the nonmeaningful condition, Hank completed *same* trials (56.12%) with significantly greater accuracy than *different* trials (48.70%), $\chi^2(1, N = 1,178) = 6.60$, $p > .05$.

Discussion

Robust differences in performance emerged across species in this analogical reasoning task. For humans, meaningfulness of stimuli had overall positive effects on their ability to reason analogically, whereas this was not uniformly the case for the chimpanzees and was completely absent in the case of the monkeys. There was also another major disconnect between the performance of human and nonhuman animals. Although meaningfulness facilitated statistically higher performance for humans, success was also observed in the RMTS task with nonmeaningful stimuli (and at a rather high overall level). Overall, the nonhuman primates not only failed to show facilitation from meaningful stimuli, but they also showed no evidence of solving the RMTS task with similarly high rates of success. It seems to be the case that numerals for monkeys and lexigrams for chimpanzees do not operate at the same conceptual level as words do for humans in terms of facilitating analogical reasoning. 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 lend both support and opposition to the previous suggestion that language-like abilities and symbolic training are integral to the ability to reason analogically. Premack (1983) concluded that language-like training is necessary on the basis of his experiments with 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 that arose between the chimpanzees urge the consideration of an alternate hypothesis.

For human participants, the known, discrete meaning of stimuli appeared to facilitate responding in the RMTS task. 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, although this difference was not statistically significant.

The representational value of each stimulus may have enabled the relational concepts of sameness and difference to be more salient to the human participants in the meaningful condition than to those who completed the task with nonmeaningful stimuli that had no inherent representational value. The discrete meaningful value of a stimulus not only enhanced its own uniqueness, but may also have removed extraneous associations it may have had to the stimulus with which it was paired.

It is especially noteworthy that some human participants failed to learn the analogical rule under any condition. On their debriefing 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 for these participants. Like children (Gentner, 1988), these adult participants may have been distracted by the surface similarities between the components of one trial (e.g., using the same elements). 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 instantaneous.

In experiments with children and adult humans, researchers have agreed 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 (Gentner et al., 1993; Rattermann & Gentner, 1998). 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 (e.g., content) similarities existed across trials. The repeated presentation of individual stimuli (albeit in various locations and combinations of pairings) could be interpreted as surface similarities that act as confounds not related to the application of analogical knowledge. If stimulus X is present in the sample pair, not only may the participant attempt to search for stimulus X among the choice pairs, but he or she may also retain knowledge of the role of stimulus X for future trials. Thus, if stimulus X is encountered again, regardless of current location or pairing, rules previously associated with it may be incorrectly applied to the current scenario.

For chimpanzees, the meaning of meaningfulness is more ambiguous. Rather than facilitating the acquisition of the analogical rule, 1 chimpanzee failed to apply the analogical rule at all unless the stimuli had discrete meaning. Even after completing the meaningful sessions first, Panzee failed to perform above chance levels during nonmeaningful sessions. This would indicate that what she learned during the meaningful sessions was not as broad an analogical rule as humans might conceive it. If that were the case, she should have transferred her knowledge across categories. 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.

Meaningfulness for Sherman had the opposite effect from that which was observed with Panzee. He completed trials above chance levels only in the nonmeaningful condition. That he could correctly complete only nonmeaningful trials may indicate that his analogical rule was confounded by the meaningfulness of the stimuli. Perhaps he attended more to the perceptual qualities of a stimulus to determine its relatedness both within the pair and between target and choice pairs. However, if using a purely perceptually bound relational matching strategy, it should follow that Sherman could succeed in the meaningful condition as well. This valuable error lends even more support that meaningfulness plays an integral role in analogical reasoning. Perhaps it is the case that meaningful stimuli add a layer of confusion to what, for Sherman, could have been a more perceptually bound task.

Our chimpanzees seemingly are not as analogical by nature as was the chimpanzee Sarah (Oden et al., 2001). The main difference between Sarah and the chimpanzees in this study is the use of special tokens or symbols for the relational concepts themselves. Although Panzee, Sherman, and Lana have been trained with lexigrams, no lexigrams specifically connote meaning of relational concepts such as *same* and *different*. Rather, they refer to concrete objects, people, and places. Sarah, however, had exactly those two tokens for sameness and difference. Perhaps to more fully apply an analogical rule, one must have the capacity to recode relational concepts symbolically.

For both Panzee and Sherman, poor performance on subsequent conditions may be indicative of the inflexibility of their application of the analogical matching rule to a categorically novel relational stimulus set. Asymmetric performance by Sherman on *different* trials in comparison to *same* trials in the condition yielding overall success (nonmeaningful) indicates that the analogical rule he may have 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 may not be as broadly conceived so that a consistent rule could be applied to *same* relations.

Both people and pigeons are predisposed to notice differences rather than similarities (Young & Wasserman, 2002). Like people and pigeons, chimpanzees exhibited some differences in performance on *same* versus *different* trials, sometimes performing better on *different* trials. If we suppose that chimpanzees are predisposed to noticing only differences, it is likely that an analogical rule was realized only through experience in matching differently related pairs with other differently related pairs. In subsequent trials, when similarly related pairs were presented, the limited analogical rule that had been learned was no longer applied.

Although conclusions about individual differences between and within species are speculative, it may be the case that the richness of the symbol systems for the animals plays an important role. For chimpanzees, lexigrams are most likely not as seamlessly integrated into the cognitive system as words, symbols, and other linguistic tools are for humans. Certainly for rhesus monkeys numerals convey some meaning, but not at a level equivalent to words for humans or even lexigrams for chimpanzees. Therefore, the inequity of meaning between the species may account for some of these disparities.

A within-species comparison may follow similar logic. There are significant differences between the usage of lexigrams by Panzee and Sherman that relate back to rearing history and early experiences. Panzee's current vocabulary includes many more lexigrams than does Sherman's (Beran, Savage-Rumbaugh, Brakke, Kelley, & Rumbaugh, 1998). Panzee uses a wider range of lexigrams and seems to have a better understanding of what they may afford her. It may be the case that Panzee's use of lexigrams is better integrated in cognitive systems that allow for analogical reasoning, thus accounting for her success with meaningful lexigrams.

The inconsistencies in the performances of our chimpanzees are also contrasted with other recent evidence that is relevant to relational matching. Vonk (2003) provided evidence that four orangutans and one gorilla without any symbol or language training succeeded in a similar non–matching-to-sample task. By making judgments between pairs of stimuli that matched on only one dimension (color or shape), Vonk claimed that the nonsymbol-trained great apes were capable of abstracting relations between relations that were not as dependent on perceptual processing. However, the apes in that study may have been dependent on an even lower level of perceptual processing. That is, those apes may have been choosing the match choice that was perceptually less similar to the sample pair. For instance, if a sample pair consisted of two items that were the same shape, the subject could have selected a choice pair in which there was only one unique shape. Likewise, if color were the critical dimension, subjects viewing a sample in which color was the same for both shapes could simply have responded to a choice pair in which there was only one color present. Following this reasoning, when stimuli can differ on only two dimensions, the judgment process may be even more tightly perceptually conceived in contrast to more complex stimuli that vary on several perceptual dimensions.

Failure by the rhesus monkeys in the current study to match relational pairs correctly in this task could be the result of one or more of a variety of reasons: The monkeys could not extract the necessary relational information from a pair of objects; relational knowledge was not encoded in such a way that it was accessible for application to novel behaviors; or perceptual properties of stimuli could not be ignored in a matching paradigm.

From their performance on previous tasks (e.g., Flemming et al., 2007), we know that monkeys can extract relational information from a pair of objects. In a two-choice discrimination paradigm, monkeys chose a pair of either identical or nonidentical objects in the presence of a discriminative cue. It may be the case that the monkeys' 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.

The relational matching paradigm used throughout this study tests analogical reasoning skills by forcing explicit expression of conceptual knowledge for identically and nonidentically related pairs. One may argue that the RMTS task may be solved using perceptual judgments of uniformity and regularity as in entropy-infused displays (i.e., Fagot et al., 2001). Perhaps it is the case that humans have a greater ability to detect small

variations in entropy (*same* pairs have an entropy value of 0, whereas *different* pairs have an entropy value of 1). If this were true, however, we would expect to find no differences between condition for humans or chimpanzees. In fact, differences between meaningful and non-meaningful conditions, regardless of which condition allows for higher levels of success, indicate that there is an inherent conceptual component to the task as it is presented in the current study. If one can detect small variations in entropy with meaningful words, one should also be able to do so with nonmeaningful combinations of letters. Furthermore, monkeys can detect these small variations in entropy (0 versus 1) but nonetheless fail the RMTS task (e.g., Flemming et al., 2007). There is no logical reason for this failure without the consideration of an analogical component to the RMTS task.

Analogy completion assumes (and relies on) conceptual knowledge of *same* and *different*. It may also be important, as evidenced by nonhuman primate performance on the task, to have concrete symbolic tokens of otherwise abstract relationships, as Oden et al. (2001) posited. Certainly, human participants in this study had a very broad conception of *same* and *different*, as humans regularly classify objects into groups on the basis of 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, Cavoto, & Cavoto, 1995; Flemming et al., 2007; Katz, Wright, & Bachevalier, 2002; Vonk, 2003; Wasserman, Frank, & Young, 2002), the acquisition of these concepts for birds and monkeys does not emerge for sometimes hundreds or thousands of trials. Human participants as young as 3 years old provide evidence that the identity–nonidentity concept emerges in significantly fewer trials (Rattermann & Gentner, 1998).

If the concepts of *same* and *different* are not as salient to nonhuman primates as they are to humans, then their accessibility when the former search for an already abstract matching rule (in RMTS) will not be readily available as a strategy. Pigeons and monkeys that have been shown to rely on arrays of multiple items to glean relational information (e.g., Fagot et al., 2001; Wasserman et al., 2002) may rely on more ecologically valid perceptually based strategies (i.e., colors and shapes) when presented with a matching-to-sample task. Unlike simple discriminations that are founded on only implicit types of conceptual knowledge, the analogical paradigm carried throughout this study requires the explicit use of the *same–different* strategy. Because the *same–different* concept is less salient and possibly more narrowly construed by nonhuman primates, it stands to reason that the application of the concept would be more difficult, if not impossible, for such an animal.

Despite the ambiguity of the exact mechanism by which analogies are realized, the present study does provide a comparison of the capacity for the acquisition of an analogical rule across three species of primates. Results from the present study reveal both common threads and disparities in the analogical reasoning skills of members of the Old World 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. Symbolic representation of categorical content, unlike relational representation, seems to have little or no facilitatory

role at all for nonhuman primates. These data may reflect one fundamental way in which nonhuman primates differ from humans in how they represent their worlds.

Acknowledgments

References

Beran MJ, Pate JL, Richardson WK, Rumbaugh DM. A chimpanzee's (*Pan troglodytes*) long-term retention of lexigrams. Animal Learning & Behavior. 2000; 28:201–207.

Beran MJ, Savage-Rumbaugh ES, Brakke KE, Kelley JW, Rumbaugh DM. Symbol comprehension and learning: A "vocabulary" test of three chimpanzees (*Pan troglodytes*). Evolution of Communication. 1998; 2:171–188.

Beran MJ, Washburn DA. Chimpanzee responding during matching-to-sample: Control by exclusion. Journal of the Experimental Analysis of Behavior. 2002; 78:497–508. [PubMed: 12507016]

Bovet D, Vauclair J. Judgment of conceptual identity in monkeys. Psychonomic Bulletin & Review. 2001; 8:470–475. [PubMed: 11700897]

Brakke KE, Savage-Rumbaugh ES. The development of language skills in bonobo and chimpanzee: I. Comprehension. Language and Communication. 1995; 15:121–148.

Clark A, Thornton C. Trading spaces: Computation, representation, and the limits of uniformed learning. Behavioral and Brain Sciences. 1997; 20:57–90. [PubMed: 10096995]

Cook RG, Cavoto KK, Cavoto BR. Same–different texture discrimination and concept learning by pigeons. Journal of Experimental Psychology: Animal Behavior Processes. 1995; 21:253–260.

Duyck W, Desmet T, Verbeke L, Brysbaert M. WordGen: A tool for word selection and non-word generation in Dutch, German, English, and French. Behavior Research Methods, Instruments, & Computers. 2004; 36:488–499.

Fagot J, Wasserman EA, Young ME. 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. 2001; 27:316–328. [PubMed: 11676083]

Flemming TM, Beran MJ, Washburn DA. Disconnect in concept learning by rhesus monkeys: Judgment of relations and relations-between-relations. Journal of Experimental Psychology: Animal Behavior Processes. 2007; 33:55–63. [PubMed: 17227195]

Gentner D. Metaphor as structure mapping: The relational shift. Child Development. 1988; 59:47–59.

Gentner D, Rattermann MJ, Forbes KD. The roles of similarity in transfer: Separating retrievability from inferential soundness. Cognitive Psychology. 1993; 25:431–467.

Gillian DJ, Premack D, Woodruff G. Reasoning in the chimpanzee: I. Analogical reasoning. Journal of Experimental Psychology: Animal Behavior Processes. 1981; 7:1–17.

Halford WK, Graeme S. Analogical reasoning and conceptual complexity in cognitive development. Human Development. 1992; 35:193–217.

Harris EH, Washburn DA. Macaques' (*Macaca mulatta*) use of numerical cues in maze trials. Animal Cognition. 2005; 8:190–199. [PubMed: 15654597]

Katz JS, Wright AA, Bachevalier J. Mechanisms of *same/different* abstract-concept learning by rhesus monkeys (*Macaca mulatta*). Journal of Experimental Psychology: Animal Behavior Processes. 2002; 28:358–368. [PubMed: 12395493]

Oden, DL.; Thompson, RKR.; Premack, D. Can an ape reason analogically? Comprehension and production of analogical problems by Sarah, a chimpanzee (*Pan troglodytes*). In: Gentner, D.; Holyoak, KJ.; Kokinov, BN., editors. The analogical mind: Perspectives from cognitive science. Cambridge, MA: MIT Press; 2001. p. 471-497.

Premack, D. Intelligence in ape and man. Hillsdale, NJ: Erlbaum; 1976.

Premack D. Animal cognition. Annual Review of Psychology. 1983; 34:351–362.

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

Rattermann MJ, Gentner D. More evidence for a relational shift in the development of analogy: Children's performance on a causal mapping task. Cognitive Development. 1998; 13:453–478.

Ross BH. This is like that: The use of earlier problems and the separation of similarity effects. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1987; 13:629–639.

Rumbaugh, DM. Language learning by a chimpanzee: The LANA Project. New York: Academic Press; 1977.

Rumbaugh, DM.; Washburn, DA. Intelligence of apes and other rational beings. New Haven, CT: Yale University Press; 2003.

Savage-Rumbaugh, ES. Ape language: From conditioned response to symbol. New York: Columbia University Press; 1986.

Savage-Rumbaugh ES, Rumbaugh DM, Smith ST, Lawson J. Reference: The linguistic essential. Science. 1980 Nov 21.210:922–925. [PubMed: 7434008]

Thompson, RKR. Natural and relational concepts in animals. In: Roitblat, HL.; Meyer, J., editors. Comparative approaches to cognitive science. Cambridge, MA: MIT Press; 1995. p. 175-224.

Thompson RKR, Oden DL. A profound disparity revisited: Perception and judgment of abstract identity relations by chimpanzees, human infants, and monkeys. Behavioral Processes. 1996; 35:149–161.

Thompson RKR, Oden DL. Categorical perception and conceptual judgments by nonhuman primates: The paleological monkey and the analogical ape. Cognitive Science. 2000; 24:363–396.

Thompson RKR, Oden DL, Boysen ST. Language-naïve chimpanzees (*Pan troglodytes*) judge relations between relations in a conceptual matching-to-sample task. Journal of Experimental Psychology: Animal Behavior Processes. 1997; 23:31–43. [PubMed: 9008861]

Vonk J. Gorilla (*Gorilla gorilla gorilla*) and orangutan (*Pongo abelii*) understanding of first- and second-order relations. Animal Cognition. 2003; 6:77–86. [PubMed: 12687418]

Washburn DA. Stroop-like effects for monkeys and humans: Processing speed or strength of association? Psychological Science. 1994; 5:375–379. [PubMed: 11539143]

Washburn DA, Rumbaugh DM. Ordinal judgments of numerical symbols by macaques (*Macaca mulatta*). Psychological Science. 1991; 2:190–193. [PubMed: 11537106]

Washburn DA, Rumbaugh DM, Richardson WK. The language research center's computerized test system for environmental enrichment and psychological assessment. Contemporary Topics. 1992; 31:11–15. [PubMed: 11538192]

Wasserman EA, Frank AJ, Young ME. Stimulus control by same-versus-different relations among multiple visual stimuli. Journal of Experimental Psychology: Animal Behavior Processes. 2002; 28:347–357. [PubMed: 12395492]

Young ME, Wasserman EA. Detecting variety: What's so special about uniformity? Journal of Experimental Psychology: General. 2002; 131:131–143. [PubMed: 11900100]
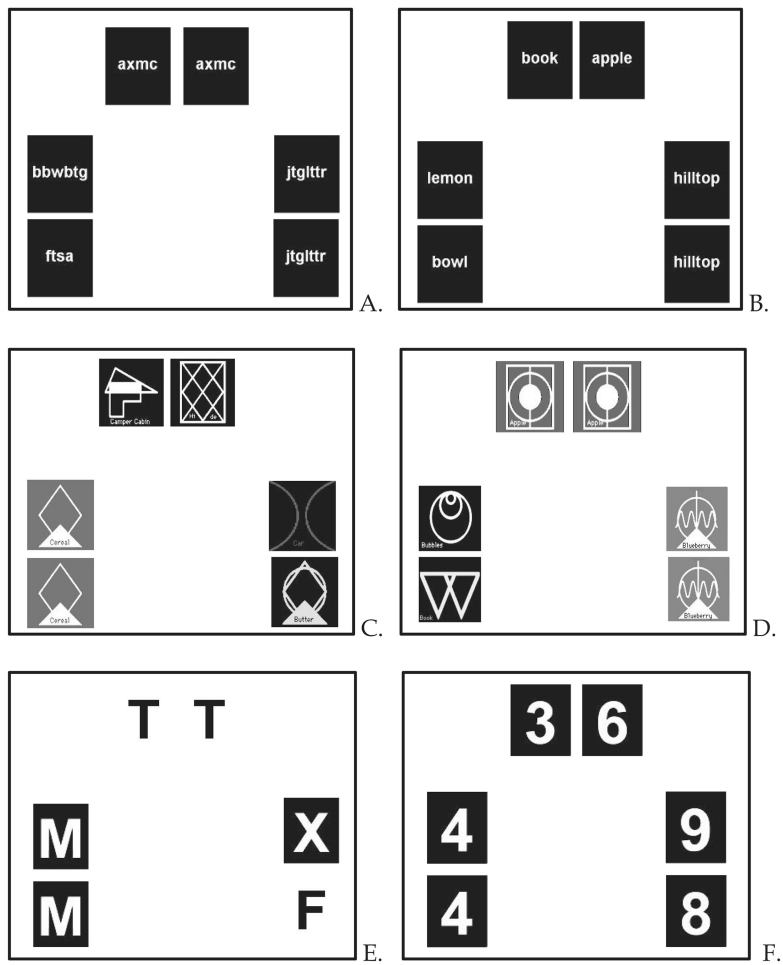
Figure 1.
Screen captures from relational matching-to-sample task for human participants, (a) same–nonmeaningful and (b) different–meaningful; chimpanzees, (c) different–nonmeaningful and (d) same–meaningful; and rhesus monkeys, (e) same–nonmeaningful and (f) different–meaningful.
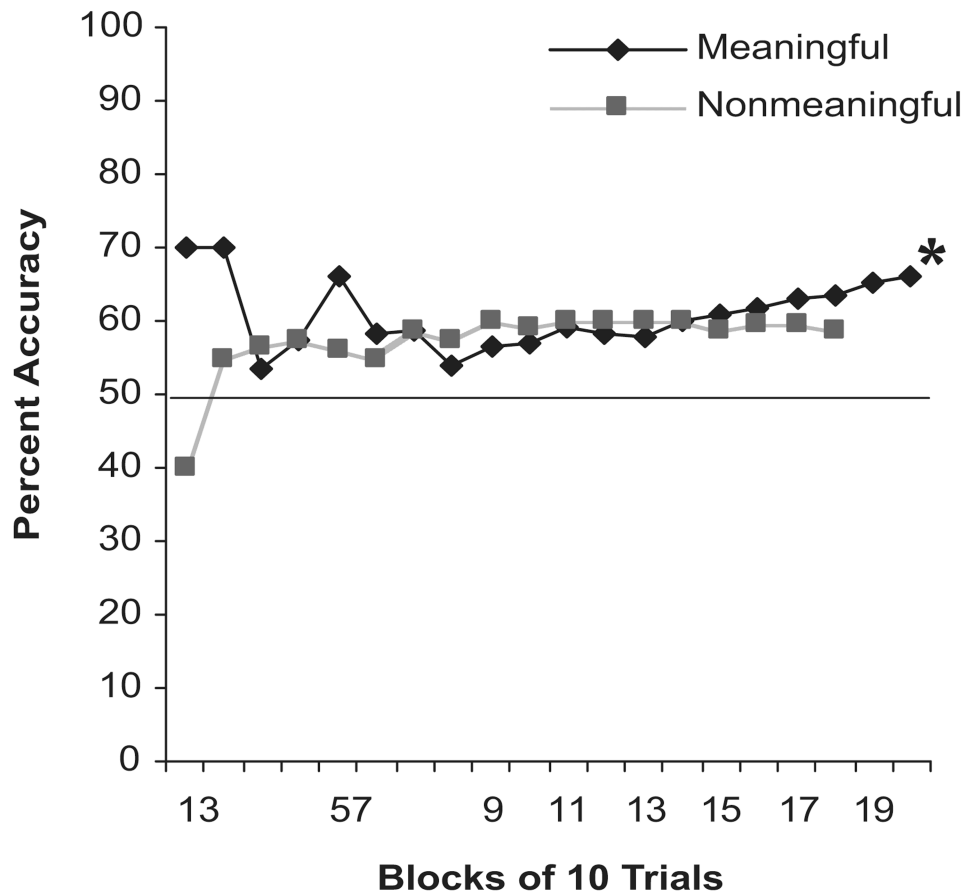
Figure 2.
Performance summary for Panzee (*Pan troglodytes*). Percentage accuracy is displayed in blocks of 10 trials for both meaningful and nonmeaningful stimuli. The horizontal line at 50% represents chance performance. Asterisks represent a performance level significantly different from chance.

Figure 3.
Performance summary for Sherman (*Pan troglodytes*). Percentage accuracy is displayed in blocks of 10 trials for both meaningful and nonmeaningful stimuli. The horizontal line at 50% represents chance performance. Asterisks represent a performance level significantly different from chance.
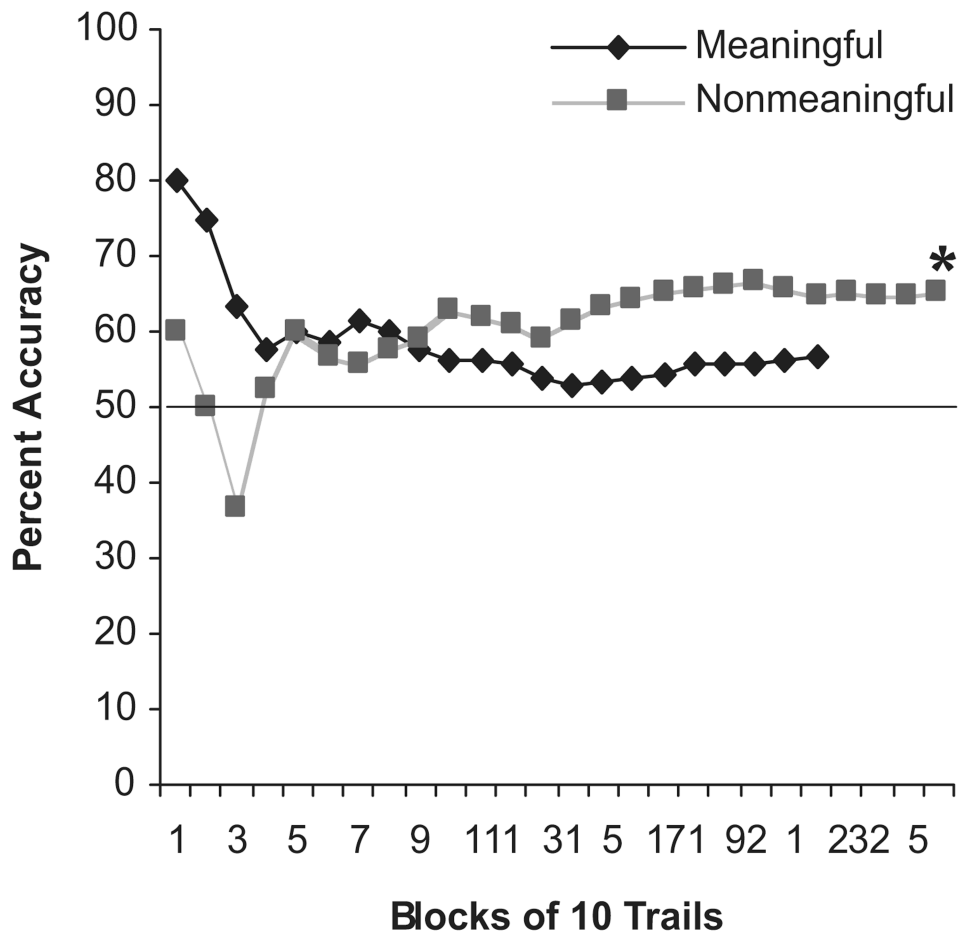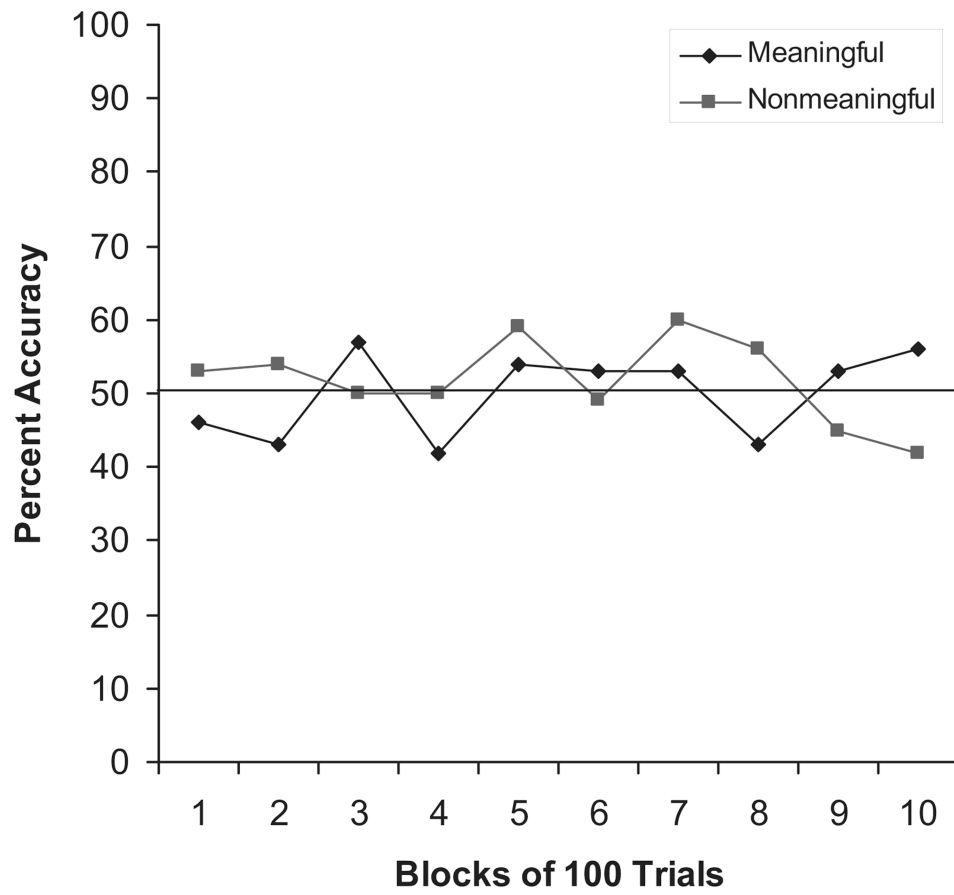

Figure 4.
Performance summary for Hank (*Macaca mulatta*). Percentage accuracy is displayed in blocks of 100 trials for both meaningful and nonmeaningful stimuli. The horizontal line at 50% represents chance performance. Performance is similar to and indicative of performance outcomes for all monkeys tested. Subsequently, graphical depictions of the nonsignificant results for all monkeys are not presented. No performance levels on any trial block were significantly different from chance.

Table 1

Performance Summaries of the Chimpanzees

Subject and condition	Order of presentation	Trials completed	% correct	z
Lana				
Meaningful	1st	238	52	0.65
Nonmeaningful	2nd	166	54	1.09
Mercury				
Meaningful	2nd	191	49	−0.21
Nonmeaningful	1st	133	56	1.47
Panzee				
Meaningful	1st	202	66	4.64[*]
Nonmeaningful	2nd	186	54	1.17
Sherman				
Meaningful	2nd	227	55	1.53
Nonmeaningful	1st	263	65	4.99[*]

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. When chimpanzees showed lasting position biases or no longer attended to the task, testing was aborted. This also affected the number of trials completed by each animal when refusal to work on the task persisted. These data represent approximately 12 sessions per animal over the span of 21 weeks of testing.

[*]
$p < .05$.

Table 2
Performance Summaries of the Rhesus Monkeys

Subject and condition	Order of presentation	Trials completed	% correct	z
Murph				
Meaningful	1st	2,766	50	0.99
Nonmeaningful	2nd	779	48	0.61
Lou				
Meaningful	2nd	1,826	51	0.94
Nonmeaningful	1st	3,599	48	−2.15[*]
Willie				
Meaningful	1st	3,242	50	0.21
Nonmeaningful	2nd	1,100	48	−0.78
Gale				
Meaningful	2nd	1,306	49	−0.44
Nonmeaningful	1st	1,174	47	−1.40
Hank				
Meaningful	1st	2,400	50	0.08
Nonmeaningful	2nd	1,178	52	1.75

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 10 sessions over the span of 2 weeks.

[*] $p < .05$.