



What do Natural Categorization Studies tell us about the Concepts of Apes and Bears?

Jennifer Vonk^{1*} and Moriah Galvan¹

¹Oakland University

*Corresponding author (Email: vonk@oakland.edu)

Citation – Vonk, J., & Galvan, M. (2014). What do natural categorization studies tell us about concepts of apes and bears? *Animal Behavior and Cognition*, 1(3), 309-330. doi: 10.12966/abc.08.07.2014

Abstract - To properly test hypotheses about the emergence of cognitive capacities in humans and other animals, it is important to test species that differ in morphology and ecology. One of the most notable aspects of animal ecology that has been championed as a factor in the evolution of cognitive skills is social complexity. However, to determine the significance of social lifestyle on an organism's cognitive abilities, it is important to test both social and less social species in similar tasks. We have conducted a series of natural categorization studies with three great ape species and recently extended this program of research to include American black bears (*Ursus Americanus*). Here we analyze the bears' choices of stimuli from our previously published categorization experiments and compare the pattern of their choices to that of apes tested previously. Like the apes, the individual bears show idiosyncratic choices that do not clearly differentiate between conceptual and perceptual strategies in performing categorization tasks. Despite their relatively asocial lifestyles, bears thus far appear to display cognitive abilities commiserate with those of the great apes. We suggest that researchers broaden their approach to study more diverse species in analogous tasks and make a concerted effort to view experimental tasks from the animals' perspectives.

Keywords – Black bears, apes, categorization, cognition

Evolutionary psychology can perhaps be best applied to comparative studies with an emphasis on testing many ecologically diverse species on similar tasks in order to shed light on factors that predict the presence (or absence) of particular traits or abilities. Categorization is a topic that received much attention in the early days of comparative cognitive research and was tightly connected to an early emphasis on perceptual processes. However researchers often treated humans and non-humans as dichotomous categories rather than focusing on the unique perceptual and cognitive capacities of different species. Initial attempts to study categorization in non-human animals were largely motivated by interest in the kinds of concepts formed by animals, particularly the depth and breadth of category membership. Primary research questions focused on the degree to which animal concepts could be described as abstract rather than concrete (Fize, Cauchoux, & Fabre-Thorpe, 2011; Herrnstein, 1990; Roberts & Mazmanian, 1988; Vonk & MacDonald, 2002; 2004; Vonk & Povinelli, 2012; Zentall, Wasserman, Lazareva, Thompson, & Ratterman, 2008). This research was considered important as part of the larger goal of determining human cognitive uniqueness. Were humans alone in the animal kingdom with regard to the ability to form abstract concepts that were not tied to observable features, such as freedom, death, mental states, and so on? And, if this was the case, what factors were responsible for the emergence of this capacity? Was it human-like language or the capacity for symbolic representation (Premack, 1983)?

Interest in the types of features used by animals to accurately perform experimental tests of categorization was of somewhat secondary importance (Jitsumori & Ohkubo, 1996). This prioritizing of research goals led to a somewhat nebulous state of affairs in the animal concept formation literature (see also Soto & Wasserman, 2010). For instance, it was clear that animals could perform categorization tasks at above-chance levels but it was not clear what this really meant for the types of concepts held by the animals. In our own work, we have shown that various ape species (orangutans, gorillas, and chimpanzees) are capable of learning category discriminations and transferring at above chance levels to novel stimuli at each level of abstraction (from concrete to abstract) in two-alternative forced-choice and match-to-sample tests, but we have not been able to conclusively determine how these apes reached significant levels of transfer or matching (Vonk, 2002; 2013; Vonk & MacDonald, 2002; 2004; Vonk, Jett, Mosteller, & Galvan, 2013). That is, it is somewhat unclear whether performance was driven by particular perceptual features of the pictorial stimuli, or was indicative of abstract overarching concepts.

In our earlier papers, we conducted analyses of the specific features of the stimuli used in our discrimination tasks largely as an effort to show that no single relevant or irrelevant feature could be used to allow positive transfer. On one hand, if relevant features gain control of responding it becomes difficult to disentangle concept-based accounts from feature or rule-based accounts of performance (see also Marsh & MacDonald, 2008). Reliance on irrelevant features, on the other hand, indicates that the subjects have not formed the overarching concept in question (D'Amato & Van Sant, 1988; Schrier, Angarella, & Povar, 1984). Most of our analyses indicated individual and esoteric patterns of response. That is, we determined that individual apes had clear preferences for selecting particular photos in our photo sets, but they did not appear to rely on the use of experimenter-predicted cues such as the presence of the eyes when categorizing animals versus non-animals or requiring a view of the entire body of the animal when categorizing primates versus non-primate animals or carnivores versus non-carnivores. Instead, we found simply that some of them generally preferred colorful objects, such as red cardinals and apples. Other researchers have also found that preference for reddish coloration may have driven the categorization of human versus non-human photos in tests with monkeys, because human photos were more likely to contain reddish patches due to clothing worn in the images (D'Amato & Van Sant, 1988). However, in our work with orangutans, orangutans were no less accurate with black and white photographs when categorizing orangutans from other apes, despite the highly predictive cue of reddish hair (Vonk & MacDonald, 2004). Thus, animals do not always use cues that are highly predictive when performing categorizations and may be attending to features that do not appear obvious to human researchers.

In designing experiments for non-humans, humans are often anthropocentric in their approach – both when framing the question and interpreting the results (Rivas & Burghardt, 2002). Often, we fail to consider the perceptual and physiological systems of the animal subjects adequately and often we assume that our own human-centric view of the world is more widely shared than perhaps is warranted. We ask animals to categorize exemplars from experimenter-defined categories based on our own experience and knowledge of the world. Often we fail to provide cues that may be prioritized by animals, such as olfactory and tactile cues. Given these failures, it is impressive that animals reach such high levels of performance in our tasks. Our own approach has been far from perfect and it is well known that laboratory studies suffer from failures of ecological validity (Vonk & Shackelford, 2012). However, the lab approach does allow us to carefully control the presence or absence of features and experiences of subjects, which should facilitate a better understanding of the thought processes of our non-verbal subjects (*cf* Soto & Wasserman, 2010). Careful post-hoc analyses of responses can also indicate the features used to perform discriminations.

Early investigations into animal categorization sometimes determined interesting patterns of responding that shed some light on the mechanisms underlying performance post-hoc. For example, several researchers have tested the ability of various primate species to categorize natural objects such as foods, humans, and members of their own and closely related species (D'Amato & Van Sant, 1988; Fujita, 1987; Jitsumori & Matsuzawa, 1991; Schrier & Brady, 1987; Schrier et al., 1984; Yoshikubo, 1985). Despite significant positive transfer, it was often unclear whether monkeys acquired the relevant “concepts” as opposed to attending to distinctive perceptual features present in only one category (*cf*

Fagot, Martin-Malivel, & Dépy, 1999). For instance, in a test of the “human” concept, D’Amato and Van Sant (1988) examined the errors made by monkeys and discovered that several monkeys incorrectly responded to a slide depicting a reddish piece of watermelon. Further inspection revealed that about thirty percent of the human slides involved patches of red, usually clothing. These findings suggest that the monkeys’ performance was controlled at least in part by a potentially irrelevant feature such as color. D’Amato and Van Sant (1988) questioned whether the monkeys were operating on the basis of true conceptual learning given that they continued to classify slides containing red patches as humans, even when they clearly belonged to the non-reinforced category. Likewise, Jitsumori and Matsuzawa (1991) found that monkeys showed some transfer to new pictures of full frontal and rear views of humans and silhouettes but did not show transfer to human faces, regardless of whether the images were ‘close-ups’ or from farther away. Both studies cast doubt on the idea that subjects formed a general concept for “human.” It is problematic for the study of concept formation when a task can be acquired and significant transfer obtained without requiring that subjects are using the same cues that predict category membership based on experimenter-defined categories. At the very least, the onus is then on the researchers to identify which features are being used, and therefore what kind of concept is really represented in these tasks.

After the initial foray into animal categorization, researchers became more methodologically skilled in assessing the features used by their subjects to perform these tasks. Advanced computer software techniques allowed researchers to isolate and manipulate various features in images to determine how such features and manipulations affected performance. However, to date, such manipulations have been performed primarily on relatively concrete level concept discriminations. For example, in testing whether monkeys could categorize images of humans, Schrier and Brady (1987) made various modifications to slides of humans, and found that monkeys were more likely to include scrambled or upside down humans in the “human” category relative to slides of silhouettes or non-human primates. As in the previous studies, such a finding made it doubtful that the monkeys had formed an overarching concept for ‘human’. Researchers have found similar patterns in work with pigeons. Jitsumori and Ohkubo (1996) manipulated orientation and found that orientation controlled responding both when pigeons categorized humans and background scenes. Pigeons could categorize stimuli on the basis of either orientation or background scenes. When these experimenters manipulated orientation (right side-up or upside-down) of the humans or backgrounds, they found that pigeons attended more strongly to the background orientation. Others have obtained what they consider to be evidence that animals form true categories through the use of techniques in which background is separated from foreground objects in training. Unlike pigeons, Range, Aust, Steurer, and Huber (2008), for example, found that domestic dogs continued to classify dog photos correctly even when the dog images were placed on previously neutral and unreinforced backgrounds. Vogels (1999) found that rhesus monkeys can learn to categorize images of trees and fish, and that responding is not controlled by simple aspects such as form, color, or texture. In addition, scrambling the images disrupted performance, suggesting, at the very least, that it is a combination of features taken together, albeit low-level features, rather than any single feature, that controls categorization. This brings us to the point that it is important to consider that features are not necessarily processed in isolation (Shimp, 2004). Researchers tend to view each stimulus feature as an independent attribute, when instead animals, as well as humans, may view features in combination or as part of a global conglomeration of features.

Other more recent experiments have demonstrated that pigeons may use axis of orientation to discriminate images of different types of animals, such as birds and mammals (Cook, Wright, & Drachman, 2013). These findings suggest that even relatively inclusive/abstract categories may be discriminated using simple, perceptual features rather than overarching concepts. Martin-Malivel and colleagues (Martin-Malivel, Mangini, Fagot, & Biederman, 2006) used a clever method to compare the information used by humans and baboons when categorizing human and baboon faces. They found that the baboons focused heavily on the eye region. In addition, by creating ambiguous human-baboon morphs, they found that only baboons used pixel similarities between the probes and training images to inform their categorization responses, suggesting a different mechanism was used by baboons and humans. The authors deemed this procedure an advance over prior methods that involved deleting particular features of the

stimuli to measure control over responding, given that such a procedure requires the experimenters to decide a priori what aspects are important to delete (as in Marsh & MacDonald, 2008; Martin-Malivel & Fagot, 2001). Researchers have also utilized artificial categories in which typicality can be manipulated; a prototype for one category contains the most features in common with other exemplars and is also the most distinctive from other categories (Jitsumori, Ohkita, & Ushitani, 2011). These authors suggest that a process by which animals, including humans, attend to general, diagnostic aspects of objects rather than to idiosyncratic aspects could underlie categorization of natural objects in many species.

Aside from focusing more heavily on the types of concepts formed, rather than on the features used to form such concepts, researchers have historically focused almost exclusively on non-human primates and pigeons when examining categorization in non-humans (Brown & Boysen, 2000; Cook, Wright, & Drachman, 2012; Herrnstein, 1979; Herrnstein, Loveland, & Cable, 1976; Kendrick, Wright, & Cook 1990; Lazareva, Freiburger, & Wasserman, 2004; Lazareva, Soto, & Wasserman, 2010; Tanaka, 2001; Zentall et al., 2008). In part this is due to the fact the primates and birds have sophisticated visual systems and the majority of categorization experiments present the subjects with visual stimuli (Soto & Wasserman, 2010). With recent years, comparative psychology has witnessed an explosion of research expanding traditional areas of research to previously unstudied or less well-studied species, such as cephalopods, cetaceans, reptiles, felines, corvids, and various other species (Vonk & Shackelford, 2012). One of the largest growth areas in comparative psychology has been the relatively recent expansion of research on canine cognition (reviewed in: Hare, 2007; Kubinyi, Virányi, & Miklósi, 2007; Miklósi & Topál 2004; Udell, Dorey, & Wynne, 2012), precipitated to a large degree by Hare's domestication hypothesis (Hare, Brown, Williamson, & Tomasello, 2002; Hare & Tomasello, 2005), in which it was proposed that the cognitive capabilities of dogs have been altered through the process of domestication to interact with humans. In particular, it is proposed that dogs have acquired keen abilities to read human communicative cues and perhaps emotion cues as well. Comparative psychology has sometimes lacked the overarching theoretical framework that guides much of the work in other areas, such as evolutionary biology. Although such theories as the domestication hypothesis have their detractors (Udell, Dorey, & Wynne, 2008; 2010; Wynne, Udell, & Lord, 2008);, it is only through developing broad theoretical frameworks and explicitly testing their hypotheses that researchers can develop a full understanding of the evolution of cognition in various species, as well as humans.

One of the other popular theories in comparative psychology has been the Social Intelligence Hypothesis, which posits that animals will have superior social cognitive skills if they have evolved to live in large, complex social groups (Dunbar, 2003; Humphrey, 1976; Jolly, 1966). Although this hypothesis has been useful and broadly applied, one of the major failings of comparative psychology today is that we have not adequately focused attention on disproving the theory. Whereas we have gathered much confirming evidence by showcasing the cognitive skills of social species such as primates, cetaceans, corvids, and canines, we have not tested relatively less social species on the same tasks or even with the same broad goals. We have also failed to adequately consider additional factors such as foraging specializations, as per the Technical or Foraging Hypothesis (Milton, 1981, 1988), which posits that challenges faced when foraging or extracting food may drive the evolution of advanced cognitive abilities.

In our lab, we have recently focused on studying the cognitive abilities of bears; in particular American black bears (*Ursus americanus*) and Grizzly bears (*Ursus arctos horribilis*). The bear family is of interest because bears are relatively non-social, but they are large-brained and have a long period of dependency as juveniles (Gittleman, 1986). Furthermore, within the bear family, foraging ecologies differ markedly and range from exclusively herbivorous (Giant Panda, *Ailuropoda melanoleuca*) to exclusively carnivorous (Polar Bear, *Ursus maritimus*), with other species, such as black and grizzly bears, being omnivorous, which represents a more generalist approach. By studying bears extensively we can evaluate the relative importance of various ecologies for influencing cognitive development. In addition, we can compare their cognitive skills to those of their fellow carnivores; canines, felines, and pinnipeds, which vary in sociality, in order to determine the true importance of social lifestyle in shaping cognitive abilities. Thus, acquiring a more nuanced understanding of the cognitive abilities of bears can assist researchers in

deciding the relative relevance or importance of social-living, foraging challenges or domestication in shaping cognitive abilities in carnivores, according to the hypotheses outlined above. Importantly, we have not yet been able to test bears on the suite of social cognitive tests that have been presented to animals such as corvids and primates. However, we have been able to utilize popular touch-screen methodology to present them with identical experiments that we presented to orangutans, gorillas, and chimpanzees, providing a direct comparison of ability and strategy. Indeed it is important for evaluating whether the benefits of group-living are limited to the development of social cognitive skills or confer more general cognitive advantages, such as with regard to problem-solving in general, for non-social species to be tested on both social and non-social tasks that provide a direct comparison to the abilities of more social species. Thus far, we have focused on assessing black bears' ability to exhibit cognitive dissonance (West, Jett, Beckman, & Vonk, 2010), make quantity comparisons (Vonk & Beran, 2012), utilize spatial memory (Zamisch & Vonk, 2012), and form natural categories (Vonk, Jett, & Mosteller, 2012), including categories depicting social relationships (Vonk & Johnson-Ulrich, 2014). In each of these tasks, with the exception of the foraging study, we have found the performance of black bears to equal or exceed that of great apes tested in similar tasks. These findings reflect the suspicions of Gordon Burghardt who conducted a series of innovative studies with bears beginning in the 1970s (Bacon 1980; Burghardt, 1975, 1992; Bacon & Burghardt, 1976a, 1976b, 1983) and found them to be quick and voracious learners.

Burghardt and colleagues also determined that bears could discriminate colors (Bacon & Burghardt, 1976b) – an important determination that makes them suitable candidates for more complex categorization tasks. These tests revealed that bears could discriminate between pairs of colors including red, green, blue, black, white, and grey. In addition, Kelling and colleagues (2006) extended this work to the giant panda, suggesting that color vision was pervasive within the bear family. Dungl, Schratter, and Huber (2008) also indicated that giant pandas could learn to discriminate two-dimensional shapes and maintain their memory for these stimuli for a long period of time. Perdue and colleagues (Perdue, Snyder, Pratte, Marr, & Maple, 2009; Perdue, Synder, Zhihe, Marr, & Maple, 2011) tested the spatial learning and memory capacities of pandas as well, helping to lay the foundation for establishing that bears embody the necessary discriminative and memory abilities for performing categorization tasks.

In our work, we presented black bears with the same two-alternative forced-choice procedure on a touch-screen computer that we had used with apes previously (Vonk & MacDonald, 2002; 2004; Vonk et al., 2013). In these studies we manipulated levels of abstraction such that the subjects had to discriminate between members of concrete level categories in which the exemplars shared many physical features (such as black bears versus humans). They also discriminated between members of intermediate level categories, such as primates, carnivores, and hoofstock, in which exemplars shared fewer features within categories and there were still notable similarities between categories. Lastly, we presented them with a relatively abstract level discrimination of animals versus non-animals, where members of the animal category were perceptually diverse, including such animals as insects, birds, reptiles, fish, and mammals. With the more abstract categories, category members were more perceptually variable but categories themselves may have also been more distinct. We found that bears rapidly learned relatively abstract category discriminations and maintained this knowledge for more than a year (Vonk et al., 2012). Here we returned to the data from our extensive categorization experiments in an attempt to determine which features (if any) were used by our bears to inform their choices, and how these strategies compared to those of apes tested previously (Vonk & MacDonald, 2002; 2004; Vonk et al., 2013). We hypothesized that bears would use perceptual cues such as color, facial features, especially eyes, and typicality of member of a category to control responding in these tasks. Thus we expected both feature-matching and prototypicality to influence responding. Based on prior findings with apes, we also anticipated individual differences in particular image preferences.

Method

Three captive adult American black bear siblings (one female and two males) were tested; Bella, Dusty, and Brutus. The bears had participated in studies of cognitive dissonance (West et al., 2010) and spatial memory (Zamisch & Vonk, 2012), but had not previously been tested in a categorization task, or a study that involved making choices on a touchscreen computer. The research, which is reported in detail in Vonk et al. (2012) took place in an off exhibit area of the bears' enclosure at the Mobile Zoo in Wilmer, AL, which consisted of two pens (3 m x 2.4 m each) separated by heavy chain link and a 1.2 m wide human access area that spanned the front of the two bear pens. Each pen had a doorway with a vertically sliding gate that allowed access to a pathway leading to the outdoor enclosure. The bears were housed with their mother who did not participate in testing due to lack of motivation. The subjects were all born in captivity and were about six years old at the beginning of testing. Testing occurred over a period of two years.

For testing, an experimenter placed a rolling cart with an embedded 19" Vartech armorall capacitive touchscreen, displaying two images, one on either half of the screen, in front the isolated bear. The experimenter stood behind the cart with a laptop that displayed the bear's response and could not direct the bears' choices. The bears were reinforced with food from the experimenter immediately following a correct response, which was cued by distinct auditory signals. The trial responses were automatically stored. The experiment was programmed in Real Basic 2006. The bears were tested on several discriminations, varying in complexity from concrete to intermediate to abstract. At the concrete level, the bears were reinforced for choosing images of conspecifics (black bears) (S+) while not choosing images of humans (S-), although this discrimination is not included in the following analyses for this article. At the concrete/intermediate level, the bears were reinforced for choosing images of polar bears (S+) versus other species of bears (S-). At the intermediate level, Bella and Brutus were reinforced for choosing images of primates (S+) while not selecting images of hoofstock (S-). Dusty, conversely, was reinforced for choosing hoofstock (S+) over primates (S-). This counterbalancing was implemented to test the idea that it is the level of category and not the particular exemplars that determine ease of category acquisition. At the intermediate/abstract level, the bears were reinforced for choosing images of carnivores (S+) while not choosing images of non-carnivores (S-). The final discrimination was at the abstract level, where the bears were reinforced for selecting images of animals (S+) instead of non-animals (S-). Other than the intermediate level discrimination, the tasks increased in inclusiveness within the same categories from concrete to abstract where each more abstract category subsumed the lower level category (e.g., black bears, polar bears/bears, carnivores, animals). The discriminations were created to mirror those presented to great apes tested previously (orangutans; Vonk & MacDonald, 2004, gorilla, Vonk & MacDonald, 2002, and chimpanzees, Vonk et al., 2013), except that the particular species represented were tailored to include the subjects and their own families and orders. Thus, apes discriminated their own species from humans at the concrete level and primates versus non-primates at the intermediate level.

The order of presentation of discriminations was counterbalanced between bears, to negate order effects that may have occurred from presenting the discriminations in a uniform manner. Thus, Bella received the discriminations in the following order: intermediate/abstract, intermediate, concrete/intermediate, concrete, and finally, abstract. Brutus received the discriminations in the following order: concrete, concrete/intermediate, abstract, intermediate/abstract, and intermediate. Lastly, Dusty proceeded through the discriminations backwards from the most abstract to the most concrete. All bears received the same S+ and S- images, in the same order in the stimulus set within a discrimination, with the exception of Dusty at the intermediate level (primate vs. hoofstock). Procedures were identical to those used with apes except for the following exceptions. The gorilla (Vonk & MacDonald, 2002) and orangutans (Vonk & MacDonald, 2004) did not receive concrete/intermediate or intermediate/intermediate discrimination tasks. In addition, they received sessions of ten rather than 20 trials and completed fewer sessions per test day, compared to the bears and the chimpanzees.

In order to determine what features the bears may have been using to inform their choices in the discriminations, we identified the images that were the most (and the least) commonly chosen by the subjects. We did this to identify whether more typical members of a category were chosen frequently or infrequently, and whether images containing unique features were selected at high or low rates. Analyzing the bears' choices in this manner allowed us to determine whether they were relying on particular features or prototypicality of exemplars to guide responding. We calculated the proportion of times an individual image was selected during presentation of that particular data set to verify which images were selected with high (75% or more of the time) and low frequency (25% or less of the time). We chose these particular cut-offs to create a clear differentiation from chance (50%) and to equate the distance of 'high' and 'low' images from chance. In addition, performance at 75% is the lowest level of performance that is significantly above chance on a single 20-trial session, so seemed a natural cut-off for images that were to be considered chosen at high rates.

Once we determined the high and low frequency images for each subject in all discriminations, two raters, who were naïve to the hypotheses of the study, were asked to respond to a set of seven questions similar to those in Roberts and Mazmanian (1988) and Vonk et al., (2013). The seven questions (seen in Table 1) inquired about different characteristics of the image, such as the prominence of the animal's eyes and limbs, the proportion of the animal's face present in the image, and the prototypicality of the animal in the image. The raters were asked to rate the image for each of the seven questions on a 5-point Likert scale ranging from 'very little' to 'very much.' They were given a clear set of written instructions and some example images and ratings. By doing so we could determine whether the features of the images chosen at high rates differed from those of images chosen at low rates. If the features across these categories (high and low) differed, then these features likely controlled responding. If the features did not differ across high and low images then they could not have controlled responding. For this reason, we analyzed the frequency of responses to both S+ and S- images. We were not interested in the correctness of the response, but the choice itself and what might have motivated it. Of course, frequency of selections must have been influenced by reinforcement, but we were interested in S+ images that were nonetheless chosen at low rates and S- images that were nonetheless chosen at high rates because these indicated choices that were not driven by reinforcement.

Table 1. Questions for categorization of images by human respondents.

Question	
(a)	How much space in the animal picture does the animal(s) take?
(b)	How colorful is the animal(s)?
(c)	How much of the animal(s)' face is shown in the picture?
(d)	How much of the entire animal(s) is shown?
(e)	How apparent are eyes in the animal picture?
(f)	How apparent are limbs in the animal picture?
(g)	How well does this picture match your ideal picture of a member of that class of animals?

Note. Raters answered questions a-g using a 1-5 Likert scale (1=very little, 5=very much) for images that were chosen at high rates (greater than 75% of the time) or low rates (lower than 25% of the time).

Results

Rater Agreement with Photographs

We correlated Rater 1 and 2's scores for each image within the discrimination categories and according to whether the photos belonged to high or low frequency categories. For the 'high' images, the correlations across all discriminations ranged from $r = 0.29$ to 0.89 , all $p < 0.01$. For 'low' images, we found that correlations across the intermediate (primate/hoofstock) and intermediate/abstract (carnivore/noncarnivore) discriminations ranged from $r = 0.32$ to 0.89 , all $p < 0.01$. However, for the

abstract discrimination (animal/nonanimal), questions a, c, d, and e yielded rater correlations that were non-significant ($r = 0.051, -0.10, 0.06, r = -0.01$, respectively). Questions b and f yielded significant correlations: $r = 0.72, 0.35, p < 0.01$. Finally, question g yielded a correlation of $r = 0.28, p < 0.05$. Rater correlations can be viewed in Table 2. A discrepancy in rater agreement for questions a, c, d, and e of the low frequency animal/non-animal discrimination presumably occurred because the images chosen at low rates were generally non-animal images (which belonged to the categories of landscapes, toys, food, and clothing). Our questions for the raters were less relevant for many of these images – thus contributing to the lack of rater agreement. Our intention was to include these questions to consider whether non-animal images that may have included face-like or eye-like features (e.g., how much does a car with headlights and a grill look like it has face-like features?), for example- may have been more often chosen at high rates than non-animal images that did not include such features. Correlations for photos chosen with high frequency from the animal/non-animal discrimination (the majority of which were animal images) were reasonable, ranging from $r = 0.58$ to $0.89, p < 0.01$, substantiating the idea that reliability was simply not appropriate for dealing with the non-animal images.

Comparison of Features of “Low” and “High” Images

Recall that for each subject, we had the raters score aspects of the images that were chosen at high rates (greater than 75% of the time) or low rates (less than 25% of the time). We took the median score for Rater 1’s ratings of the photographs on each question listed in Table 2 and used Wilcoxon signed ranks tests to compare the scores on each question for the frequently chosen versus the infrequently chosen images. We found that the features of the frequently chosen images did not differ significantly from the features of the infrequently chosen images according to the raters’ median scores for any of the subjects (all $p > 0.08$). This result indicates that the bears were not aided or hindered by such features as size of the image, amount of the body visible in the image etc. Of course it is possible that the bears were attending to features not identified as important by the researchers, but it is also possible that they used overarching concepts to categorize the stimuli.

Category of Choices

In order to determine if the subjects had preferences for a specific group of animal, we conducted analyses similar to those of Roberts and Mazmanian (1988). For each discrimination we classified the images by the most descriptive taxonomic biological classification that encompassed the maximum possible images for the set. For example, from the discriminations animal/non-animal and carnivore/non-carnivore, animal images and non-carnivore images were classified by biological class (e.g., bird, insect, reptile, etc.) For mixed bear (from the polar bear/mixed bear discrimination), classification was by biological species. Finally, for the carnivore/non-carnivore discrimination, classification was done by the biological family. We analyzed the percentage of image choices from a particular taxonomic group across all sessions for each set of photos within a discrimination. This information is presented in Table 3. Note that some of our analyses involved S- and some involved S+ stimuli. We conducted Friedman tests to compare the rate of selection of images of a particular taxonomic group to determine if the bears were selecting animal groups at similar rates. We found that there were significant differences between bears in each category of stimuli. We did not analyze choices in the most concrete level discrimination because species did not vary in black bear or human categories. For the same reason we did not analyze the polar bear set.

Table 2

Median scores of Rater 1 for 'high' and 'low' stimuli

	Question	High Median	SD	r	Low Median	SD	r
Intermediate (Primate/Hoofstock)	Size of animal	3.67	0.58	0.81**	3.67	0.58	0.70**
	Colorful	2.00	0.00	0.67**	2.00	0.00	0.62**
	Face	4.67	0.58	0.83**	4.67	0.58	0.83**
	Entire Body	3.00	0.00	0.84**	3.00	0.00	0.87**
	Eyes	4.33	1.15	0.51**	4.33	1.15	0.61**
	Limbs	2.00	0.00	0.79**	2.33	0.58	0.80**
Intermediate/Abstract (Carnivore/NonCarnivore)	Size of animal	3.67	0.58	0.84**	3.00	0.00	0.68**
	Colorful	2.00	0.00	0.66**	2.00	0.00	0.82**
	Face	4.00	0.00	0.66**	3.00	0.00	0.69**
	Entire Body	3.00	0.00	0.86**	4.00	0.00	0.89**
	Eyes	4.00	0.00	0.63**	2.00	0.00	0.81**
	Limbs	1.67	0.29	0.86**	3.00	0.00	0.67**
	Prototypical	3.33	0.58	0.29**	3.33	0.58	0.57**

Table 2(cont.)

	Question	High Median	SD	r	Low Median	SD	r
Abstract	Size of animal	3.33	0.58	0.66**	3.00	0.00	0.05
(Animal/NonAnimal)	Colorful	2.00	0.00	0.79**	2.33	0.58	0.72**
	Face	3.67	0.58	0.65**	2.67	0.58	-0.10
	Entire Body	3.67	0.58	0.89**	3.67	0.58	0.06
	Eyes	3.33	0.58	0.60**	1.33	0.58	-0.01
	Limbs	2.67	0.58	0.80**	1.00	0.00	0.35**
	Prototypical	3.67	0.29	0.58**	3.33	0.58	0.28*

*-significant at $p < 0.05$ ** -significant at $p < 0.01$

Note. R = Pearson r for rater correlations.

Table 3

Means and Standard Deviations for Proportion of Choices of Images of Different species (or Object Types) Within Each Discriminations for Each Subject

Discrimination	Stimulus Species	Brutus (SD)	n	Dusty (SD)	n	Bella (SD)	n
Polar Bear	Polar Bear	59.53 (17.92)	107	73.03 (12.09)	50	65.55 (19.25)	75
Mixed Bear	Black Bear	24.85 (32.06)	66	34.09 (39.14)	42	31.99 (39.26)	68
	Panda	35.56 (32.40)	90	26.22 (33.21)	41	48.88 (42.64)	67
	Brown	39.97 (24.65)	107	28.04 (18.31)	50	33.02 (21.44)	75
	Spectacled	43.61 (31.28)	107	24.00 (30.71)	50	37.11 (33.29)	75
	Sun	42.042 (49.74)	107	25.50 (19.50)	50	30.69 (29.33)	75
	Sloth	42.99 (49.74)	107	10.00 (26.73)	50	35.33 (45.60)	75
Primate	Chimpanzee	62.50 (48.80)	64	22.45 (42.16)	49	72.55 (45.07)	51
	Gorilla	57.81 (49.78)	64	59.18 (49.66)	49	80.39 (40.10)	51
	Orangutan	63.33 (48.60)	60	22.86 (42.60)	35	80.56 (40.14)	36
	Human	73.33 (44.60)	60	42.86 (50.21)	35	61.11 (49.44)	36
	Gibbon	84.17 (35.04)	60	27.14 (44.30)	35	88.89 (27.02)	36
	Prosimian	57.29 (49.50)	64	42.52 (46.03)	49	62.42 (46.05)	51
	New world monkey	56.58 (19.74)	64	34.48 (28.38)	49	74.93 (14.84)	51
	Old world monkey	78.71 (19.29)	64	35.29 (24.06)	49	79.92 (17.18)	51
Hoofstock	Hoofstock	31.04 (13.47)	64	65.03 (21.16)	49	23.07 (9.37)	51

Table 3 (cont.)

Discrimination	Stimulus Species	Brutus (SD)	n	Dusty (SD)	n	Bella (SD)	n
Carnivore	Ursidae	66.16 (31.81)	99	77.00 (33.82)	50	45.34 (32.68)	118
	Mustelidae	74.11 (18.04)	99	87.13 (15.16)	50	72.66 (26.70)	118
	Procyonidea	65.15 (30.96)	99	83.67 (25.31)	50	70.62 (29.45)	118
	Hyaenidae	95.25 (28.80)	99	73.00 (35.30)	50	68.64 (38.56)	118
	Felidae	74.65 (35.29)	99	53.40 (41.73)	50	67.37 (38.21)	118
	Canidae	75.25 (29.96)	99	59.33 (30.89)	50	77.54 (27.60)	118
	Viverridae	73.23 (32.80)	99	85.33 (22.24)	50	77.54 (30.22)	118
	Herpestidae	78.87 (31.24)	71	72.37 (27.75)	50	63.37 (34.66)	118
Non-carnivore	Insect	18.94 (18.40)	99	5.50 (10.77)	50	33.26 (29.89)	118
	Fish	27.44 (30.71)	99	11.33 (20.88)	50	29.38 (33.10)	118
	Bird	12.12 (24.95)	99	6.67 (19.05)	50	32.20 (37.41)	118
	Amphibian	11.11 (30.77)	99	13.00 (31.64)	50	47.46 (47.52)	118
	Mammal	38.48 (16.97)	99	42.03 (19.31)	50	35.46 (20.44)	118
	Primate	36.11 (45.61)	90	43.75 (44.13)	40	26.25 (41.32)	80
	Hoofstock	30.30 (36.99)	99	17.00 (27.87)	50	38.56 (41.04)	118
	Reptile	17.17 (33.63)	99	8.00 (23.39)	50	22.03 (34.93)	118

Table 3(cont.)

Discrimination	Stimulus Species	Brutus (SD)	n	Dusty (SD)	n	Bella (SD)	n
Animal	Insect	62.66 (34.03)	129	46.30 (35.16)	45	50.39 (33.82)	86
	Fish	69.64 (26.88)	129	70.00 (26.02)	45	77.52 (27.15)	86
	Bird	68.86 (25.21)	129	61.48 (29.48)	45	62.40 (29.82)	86
	Amphibian	76.58 (42.54)	111	88.89 (31.87)	36	67.80 (47.13)	59
	Mammal	62.86 (30.06)	129	74.44 (22.99)	45	75.78 (25.65)	86
	Carnivore	67.44 (41.41)	129	93.89 (14.26)	45	81.98 (34.22)	86
	Hoofstock	66.95 (27.21)	129	79.37 (22.23)	45	82.87 (21.86)	86
	Reptile	64.73 (29.73)	129	72.59 (32.79)	45	64.34 (33.55)	86
Non-animal	Landscape	30.12 (19.66)	86	29.89 (18.97)	86	31.07 (15.71)	86
	Toys	30.23 (26.74)	86	27.78 (21.90)	86	25.81 (24.56)	86
	Food	35.71 (21.33)	86	25.07 (16.46)	86	22.44 (18.90)	86
	Clothing	38.76 (36.75)	86	19.26 (24.61)	86	25.00 (32.21)	86

Note. Highlighted cells indicate S- categories, while non-highlighted cells indicate S+ categories. “n” indicates the total number of opportunities to select a particular category of image.

Polar bear/mixed bear. Friedman tests revealed that each bear chose images from the category “mixed bear” at different rates depending on the species of the animal depicted in the image (e.g., black bear, brown bear, panda bear, sun bear, sloth bear, spectacled bear); Brutus ($N = 49$, $\chi^2_7 = 51.79$, $p < 0.001$), Dusty ($N = 33$, $\chi^2_7 = 73.39$, $p < 0.001$), and Bella ($N = 60$, $\chi^2_7 = 34.52$, $p < 0.001$). Brutus chose the spectacled bear most often, and the panda bear least often. Dusty chose the sloth bear least often, and the brown bear most often. Bella selected the sun bear at the lowest rates, while selecting the panda bear at the highest rates. Thus, rates of selection of different animals varied widely between the subjects. Giant Pandas are probably the least typical of the bears in appearance so it is interesting that Brutus chose images of pandas at low rates whereas Bella chose them at high rates. It is possible that distinctiveness of an item within a category makes an item stand out, thus leading it to be selected more frequently (see also Beran, 2011; Vonk & Mosteller, 2013). Avoiding distinctive exemplars, on the other hand, may indicate the use of typicality to dictate choices.

Primate/hoofstock. Friedman tests revealed that each bear chose images from the category “primate” at different rates depending on the genus or species of the animal depicted in the image (e.g., chimpanzee, gorilla, orangutan, human, gibbon, prosimian, new world monkey, old world monkey); Brutus ($N = 60$, $\chi^2_7 = 39.16$, $p < 0.001$), Dusty ($N = 35$, $\chi^2_7 = 35.01$, $p < 0.001$), and Bella ($N = 36$, $\chi^2_7 = 47.21$, $p < 0.001$). As with the other bear discrimination, rates of selection of different animals varied between the subjects (see Table 3). Brutus chose gibbons and old world monkeys the most often, although he would have seen significantly more pictures of humans prior to this test. Dusty chose gorillas the most often but chose orangutans, chimpanzees, and gibbons the least often. One might argue that gorillas look the most similar to black bears, which he had not yet been tested with, but perhaps he preferred images of animals most similar to his own group. Clearly he was not just choosing images of animals with black hair, however, given that he did not chose chimpanzees and bonobos at equally high rates. Bella selected gorillas and gibbons at the highest rates. She selected prosimians at the lowest rate, which makes some sense as they probably appear the most unlike the other members of the category. She also selected humans at low rates, for which one could make a similar argument.

Carnivore/non-carnivore. Friedman tests revealed that each bear chose images from the category “Carnivore” at different rates depending on the family of the animal depicted in the image (e.g., ursidae, mustelidae, procynoidea, hyaenidae, felidae, canidae, viverridae, herpestidae); Brutus ($N = 71$, $\chi^2_7 = 42.32$, $p < 0.001$), Dusty ($N = 38$, $\chi^2_7 = 26.39$, $p < 0.001$), and Bella ($N = 86$, $\chi^2_7 = 87.27$, $p < 0.001$). Rates of selection of different animals varied between the subjects (see Table 3). Surprisingly, Brutus chose hyaenidae, felidae and canidae the most often, and ursidae the least often. Dusty chose mustelidae the most often, and canidae the least often. Bella selected ursidae at the lowest rates as well, while selecting hyenidae and canidae at the highest rates. Because bears belong to ursidae and they are both the subject’s own group, as well as a category of images that had been presented and reinforced previously, the lower frequency of choosing bears for Bella and Brutus is surprising. At the very least, the result indicates that bears do not simply select categories of stimuli that have been presented previously, or previously reinforced. Perhaps they were drawn to more novel images. In this case Brutus and Bella’s choices were quite similar but difficult to interpret. Whereas Brutus and Bella chose canidae at high rates, Dusty chose them at low rates, revealing again inconsistencies across subjects.

Friedman tests revealed that each bear chose images from the category “non-carnivore” at different rates depending on the class of the animals depicted in the image (e.g., insect, fish, amphibian, reptile, bird, mammal, primate, hoofstock); Brutus ($N = 90$, $\chi^2_7 = 115.92$, $p < 0.001$), Dusty ($N = 40$, $\chi^2_7 = 94.03$, $p < 0.001$), and Bella ($N = 80$, $\chi^2_7 = 31.49$, $p < 0.002$). As with the prior tests, rates of selection of different images varied between the subjects. See Table 3. Brutus chose mammals most often and then primates. Recall that Brutus had not yet been tested on primates versus hoofstock at the time of this test so familiarity with the species depicted in the images cannot explain these results. Rather, it appears that he generally selected mammal photos at high rates, probably because most of the prior images he had been shown were of mammals (e.g., bears, humans). Dusty chose mammals and primates at high rates as well, but chose all other categories at low rates, including hoofstock. Dusty had completed only the

animal/non-animal test before the carnivore/non-carnivore test so prior presentation or reinforcement history was unlikely to influence his results. Bella inexplicably chose amphibians most often, yet chose reptiles least often. Bella did not choose animals from categories she had been presented with previously at higher rates.

Animal/non-animal. Friedman tests revealed that each bear chose images from the category “animal” at different rates depending on the broad class of the animal depicted in the image (e.g., insect, fish, amphibian, reptile, bird, mammal carnivore, hoofstock); Brutus ($N = 111$, $\chi^2_7 = 26.34$, $p < .001$), Dusty ($N = 36$, $\chi^2_7 = 65.11$, $p < 0.001$), and Bella ($N = 59$, $\chi^2_7 = 41.12$, $p < 0.001$). As with the apes tested before, and the previous tests, rates of selection of different animals varied between the subjects (see Table 3). Surprisingly, Brutus chose amphibians and fish the most often, and mammals and insects the least often. Dusty chose insects and birds the least often, and amphibians and carnivores the most often. This result cannot be explained by Dusty having been reinforced for selecting images of carnivores because he completed the animal/non-animal test prior to the carnivore/non-carnivore or black bear/human tests. Bella selected insects and birds at the lowest rates, while selecting carnivores and hoofstock at the highest rates. In Bella’s case, she had been previously reinforced for selecting carnivores but she had not been reinforced for selecting hoofstock. As with the chimpanzees in Vonk et al. (2013), she may have been more prone to selecting images of animals that were more familiar through prior presentation although not necessarily reinforced. Note that all of the specific images at each level of abstraction were novel. That all bears selected insects at low rates is consistent with the use of features such as faces and eyes to determine membership to the animal category, although presence of the face or eyes was not present more often in photos they selected often compared to those they rarely selected (see above). However, it is the case that insects are perceptually distinct from other animals in a number of ways that may have made the bears less likely to classify them as animals. Such a finding suggests the use of perceptual features rather than a conceptual understanding of what it is to be animate in the formation of the ‘animal’ concept.

Friedman tests revealed that each bear chose images from the category “non-animal” at different rates depending on the broad categories of the objects depicted in the image (e.g., landscape, toy, food, clothing); Brutus ($N = 129$, $\chi^2_7 = 8.27$, $p = 0.04$), Dusty ($N = 45$, $\chi^2_7 = 9.13$, $p = 0.03$), and Bella ($N = 86$, $\chi^2_7 = 10.34$, $p = 0.02$). As with the previous tests, rates of selection of different images varied between the subjects. See Table 3. Brutus chose clothing most often and chose food least often. Dusty chose clothing least often, and landscapes most often. Bella also chose landscapes most often, but chose food least often. They may have chosen landscapes more often because many of the previously reinforced animal photos appeared on similar backgrounds. Again there was little consistency between individuals.

Discussion

As with the apes tested previously (Vonk & MacDonald, 2002; 2004; Vonk et al., 2013), the black bears revealed individual idiosyncratic preferences for particular photographs that were not tied to prior reinforcement history. It was not possible to directly compare the choices of stimuli to that of the apes for each level of abstraction given that the images presented to each species were selected specifically to include members of the subjects’ own species/genus/family and order. However both apes and bears were presented with images of primates, carnivores, hoofstock, animals, and non-animals. The chimpanzees additionally received the non-reinforced category of “non-orangutan apes” while the bears received the non-reinforced category of “non-polar bear bears” at the concrete/intermediate level. In addition the gorilla and orangutans were not presented with the concrete/intermediate or intermediate/intermediate tasks. However, we can compare their performance based on the level of abstraction presented and this data is presented in Tables 4 and 5. These results reveal similar levels of transfer performance across species, although orangutans and the only gorilla tested tended to acquire the concepts more rapidly in general. That overall performance is similar across such distantly related species that have evolved quite different lifestyles, but more similar diets, hints at the fact that technical, rather

than social, challenges may be more critical to the development of skills such as concept learning and problem-solving. Although, tests of social concepts in less social species, such as bears, will be an important future test as to whether group-living confers specific advantages with regard to social cognitive skills.

We can also assess in general whether the subjects were similarly or differentially affected by such features as stimulus presentation and reinforcement history, size of the animal in the image, orientation, presence of eyes and facial features, color, and so on. We found that the subjects were not generally affected by these features in any way that consistently predicted their choices in any of the discrimination tasks, based on the fact that raters did not find images chosen at high or low rates to differ on these attributes. However, the chimpanzees were most likely to be affected by prior presentation of the species depicted in the images – being inclined to select images of species that had been presented in prior discriminations, even if not reinforced (Vonk et al., 2013). In contrast, orangutans, when reinforced for selecting photos of “primates”, were not more likely to select images of species that had been shown and reinforced previously compared to images of novel species (Vonk & MacDonald, 2004). If anything, bears appeared highly unlikely to make selections based on prior reinforcement history. They also didn’t appear more likely to choose images that might have represented preferable real life objects. For instance, apes were most likely to choose foods in the non-animal category, but this was not the case for the bears. At the very least this result suggests that bears’ performance is not driven by simple stimulus-response associations.

Orangutans selected photos where the animal in the image was larger (Vonk & MacDonald, 2004). For selecting orangutans from other primates, orangutans preferred images that showed only the head of the animal. However, when presented with control primate photos that did not show faces or showed only faces, subjects performed just as well as they had when the entire head and body of the animals were shown in the photographs. Some of the photos that were incorrectly chosen at high rates by the orangutans were those that shared physical features with correct photographs. For instance, they chose a uakari when reinforced for choosing orangutans and a mouse when reinforced for choosing primates. Uakari monkeys, like orangutans, have shaggy, reddish hair. Mice may look quite similar to prosimians, which were also included in the primate photo set. Prior reinforcement or presentation history did not affect the orangutans’ performance on any of the discriminations. Orangutans also did not appear to preferentially select exemplars of categories that were more typical or more similar to images previously shown. For example, when reinforced for selecting photos of the broad category “animal”, they were just as likely to choose reptiles or fish as they were to select mammals and no less accurate with images of animals such as butterflies or ‘worm lizards.’ Bears, on the other hand, were less likely to select less typical animals such as insects. These results suggest that orangutans were more likely to use features than assessments of typicality.

Similar to the orangutans, the gorilla tested previously was not influenced by prior presentation or reinforcement history of particular images (Vonk & MacDonald, 2002). She was not distracted by intentionally confusing images such as a “horse and rider” statue or an image of clay elephants in the “non-animal” category. Like the orangutans, she was equally proficient at selecting animal photos that were atypical, such as a photo of a worm that appeared stick-like. Chimpanzees were interestingly more likely to choose images rated lower on typicality. In general, none of the apes used experimenter-defined typicality in order to inform their choices. It appears that orangutans may have been more likely to use relevant features, whereas chimpanzees used prior experimental history and bears may have used typicality to some extent. Future work should more explicitly test such differences using stimuli that are ecologically relevant to the species tested.

Table 4

Average Number of Trials to Reach Criterion Across Photo Sets by Discrimination and Subject.

Discrimination	American Black Bear			Chimpanzee		Gorilla		Orangutan				
	Brutus	Dusty	Bella	Joe	Bonzo	Zuri	DinDing	Dinar	Molek	Abby	Jahe	Puppe
Concrete	400.00	160.00	386.60	233.40	660.00	80.00	23.30	46.70	60.00	220.00	46.70	90.00
Concrete/Intermediate	560.00	250.00	380.00	526.60	780.00	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Intermediate	426.60	326.60	340.00	293.40	2186.60	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Intermediate/Abstract	660.00	333.40	786.60	566.60	1353.40	126.70	53.30	86.70	33.30	86.70	53.30	56.70
Abstract	840.00	300.00	573.40	1220.00	673.40	45.00	150.00	97.50	47.50	52.50	80.00	42.50

Table 5

Average Transfer Performance Across Photo Sets by Discrimination and Subject.

Discrimination	American Black Bear			Chimpanzee		Gorilla		Orangutan				
	Brutus	Dusty	Bella	Joe	Bonzo	Zuri	DinDing	Dinar	Molek	Abby	Jahe	Puppe
Concrete	67.50	70.00	67.50	70.00	60.00	90.00	87.50	66.68	73.90	62.50	72.30	69.60
Concrete/Intermediate	58.33	73.33	75.00	65.00	62.50	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Intermediate	72.50	72.50	75.00	70.00	35.00	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Intermediate/Abstract	70.00	75.00	72.50	62.50	65.00	64.00	72.50	55.00	82.50	71.25	75.00	63.44
Abstract	50.00	77.50	72.50	62.50	62.50	86.33	56.67	67.58	73.33	76.67	68.33	78.33

Members of all species were able to acquire concepts and show significant transfer at each level of abstraction, although the gorilla encountered the most difficulty with the intermediate level of abstraction and the chimpanzee with the abstract level of abstraction (see Table 5). It is difficult to draw conclusions regarding the effect of level of abstraction on bear categorization given that the bears were tested in different test orders. The bears that performed the best on the most abstract levels were those that encountered these tasks earlier in the set, suggesting that order of presentation may affect the ability to conceptualize abstract concepts (see Vonk et al., 2012). Larger numbers of subjects tested on similar problems with varying task experiences will be necessary to further probe this interesting issue. In the majority of experimental research programs with laboratory animals, animals are trained for an extended period of time on tasks requiring them to determine perceptual similarity before being transferred to more abstract conceptual discriminations. It is possible that this habit over-trains animals to attend to perceptual similarity causing them to overlook more conceptual category membership. It is important to note that all of the participants discussed here were experimentally naive when they began these tests, making them unique among subjects typically tested in laboratory categorization studies.

One of the most notable findings from our program of research is the ability to compare task acquisition and levels of transfer performance on identical tasks with identical training regimes across such diverse species as chimpanzees and bears. We found that bears required a comparable number of sessions to reach criterion in general, and often showed better transfer with novel photographs compared to chimpanzees. Unfortunately, it is more difficult to compare performance directly to that of orangutans or the single gorilla tested previously because of differences in training, such as number of exemplars/trials per discrimination and differences in criterion (two versus four consecutive sessions at 80% or more). In addition, test order was held constant for orangutans and the gorilla while the bears and the two chimpanzees were tested on the discriminations in random orders in order to negate the concern that later discriminations are learned better simply because of testing experience. We can compare overall patterns nonetheless, and we found that orangutans excelled at intermediate level category discriminations, which were found to be difficult for bears and a gorilla in terms of sessions required to reach criterion. Both chimpanzees and orangutans found the abstract discriminations more difficult to acquire. This was not true of bears or the gorilla. Thus, a tentative conclusion is that bears and gorillas are aided in categorization tasks when there is a significant amount of variability between categories, but are not hindered when there is also a significant amount within categories (e.g., abstract level category discriminations). Chimpanzees may do best when there is less variability within categories (concrete level discriminations). However, we must be cautious in drawing conclusions about species differences when dealing with such small sample sizes.

Lazareva et al. (2010) suggested that abstract concepts may be easily acquired because, although the exemplars within categories are perceptually diverse, there is also less overlap between categories. The category discrimination tasks traditionally presented to pigeons and other animals may therefore allow even the more abstract categories to be discriminated through the use of comparison of perceptual features (Vonk & Povinelli, 2012). It is likely that the degree of feature overlap both within and between categories is important for category discrimination in both non-humans and children (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). Intermediate level categories that contain many diverse exemplars within the category but also contain exemplars that share many features between categories (e.g., dogs versus cats), may be more difficult to discriminate than abstract level categories, such as animals versus foods, which include perceptually variable members, but that share fewer features between categories. An animal's pattern of responses with various category discriminations can therefore be informative with regard to the extent that the animal relies on between or within category similarity.

Like our gorilla, Roberts and Mazmanian (1988) found that pigeons and monkeys encountered the most difficulty with an intermediate level of abstraction. That is, despite the prediction that animals would encounter the greatest difficulty at the most abstract level, both pigeons and squirrel monkeys could learn to accurately categorize animals and non-animals at the most abstract level, but failed to acquire the bird versus non-bird discrimination at the intermediate level discrimination even with additional training (Roberts & Mazmanian, 1988). However, Kendrick et al. (1990) found that pigeons

could learn an intermediate level discrimination of birds versus mammals, if a sufficient number of exemplars (35) were presented during training. Because orangutans, gorillas, and bears showed positive transfer to novel images on the most abstract problems, it seems that these animals rely on between category dissimilarity to perform the tasks. Such a conclusion is also consistent with prior findings by Vauclair and Fagot (1996) and DePutte, Pelletier, and Barbe (2001). It may be the degree of shared features within and between the particular sets of images used that determine acquisition of categories, rather than the conceptual level of experimenter-defined categories. Others have argued that both typicality within a category and distinctiveness from other categories are related to prototypicality in natural categories (Jitsumori, et al., 2011; Rosch, 1978; Rosch & Mervis, 1975). It is important to note that the use of typicality, features, and concepts are not mutually exclusive; however, true concepts can be separated from the presence or absence of any particular feature. For instance, a human adult would recognize that a bird with broken wings that can no longer fly is still a bird.

Marsh and MacDonald (2008) examined the features used by orangutans to perform concrete level discriminations and indicated that orangutans relied upon coloration and the presence of eyes to discriminate gorillas or orangutans from other apes. Because these features are not independent of the overarching category structure, true concept formation could not be ruled out in their study. We obtained no evidence that such features controlled performance in any of the species we tested, and we additionally showed that orangutans, along with gorillas, chimpanzees, and black bears, could form more abstract concepts. The factors underlying more abstract level concept formation have yet to be identified. What our admittedly low-level analysis of performance on these tasks indicates most strongly is that much further work is needed with these and other species in order to answer important questions about the underlying mechanism of animal categorization. Future work should also pursue the question of whether animals actually perceive two dimensional images as representations of real objects (Deputte et al., 2001). Prior work addressing this issue has called this assumption into question (Martin-Malivel, 1998; Parron, Call, & Fagot, 2008; although see also Bovet & Vauclair, 1988; Fagot et al., 1999; Watanabe, 1993), potentially rendering categorization experiments meaningless except to the extent that animals can categorize perceptual patterns. Others have suggested, following sophisticated neural techniques, that the categorization of images is quite similar in monkeys and in humans (Fize et al., 2011; Kromrey, Maestri, Hauffen, Bart, & Hegdé, 2010) begging the question of whether similarities end at perceptual processes or overlap into conceptual categories. Studies of categorization have lagged behind research in other areas that are tapping into increasingly abstract constructs such as mental states. It is our hope that categorization studies will continue to be methodological trend-setters while venturing into ever more abstract territories. We hope that demonstrating the possibility of using similar methodology and stimuli between animals as different as bears and apes will inspire others to adopt a truly comparative approach to answer questions about why and how various cognitive abilities have evolved.

Acknowledgements

We are indebted to the Mobile Zoo, especially its director, John Hightower. Without his support and assistance the bear experiments could not have been conducted. In addition, special thanks to Dr. Joan Sinnott for her support, as well as the assistance of Stephanie Jett during data collection, and Zoe Johnson-Ulrich and Jennifer Hamilton for coding images used in this study.

References

- Bacon, E. S. (1980). Curiosity in the American black bear, *Ursus*, 4, 153-157.
- Bacon, E. S., & Burghardt, G. M. (1976a). Ingestive behaviors of the American black bear. *Ursus*, 3, 13-25.
- Bacon, E. S., & Burghardt, G. M. (1976b). Learning and color discrimination in the American black bear. *Ursus*, 3, 27-36.
- Bacon, E. S., & Burghardt, G. M. (1983). Food preferences in the American black bear: An experimental approach. *Ursus*, 5, 102-105.
- Beran, M. J. (2011). Chimpanzees (*Pan troglodytes*) show the isolation effect during serial list

- recognition memory tests. *Animal Cognition*, 14, 637-645.
- Bovet, D., & Vauclair, J. (1998). Functional categorization of objects and of their pictures in baboons (*Papio anubis*). *Learning and Motivation*, 29, 309-322.
- Brown, D. A., & Boysen, S. T. (2000). Spontaneous discrimination of natural stimuli by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 114, 392-400. doi:10.1037/0735-7036.114.4.392
- Burghardt, G. M. (1975). Behavioral research on common animals in small zoos. In Institute of Laboratory Animal Resources (Ed). *Research in Zoos and Aquariums: A Symposium Held at the Forty-ninth Conference of the American Association of Zoological Parks and Aquariums* (pp. 103–133). Washington, D.C.: National Academy of Sciences.
- Burghardt, G. M. (1992). Human–bear bonding in research on black bear behavior. In H. Davis & D. Balfour (Eds): *The Inevitable Bond* (pp. 365–382). Cambridge, UK: Cambridge University Press.
- Cook, R. G., Wright, A. A., & Drachman, E. E. (2013). Categorization of birds, mammals, and chimeras by pigeons. *Behavioral Processes*, 93, 98-110. doi: org/10.1016/j.beproc.2012.11.006
- D'Amato, M. R., & Van Sant, P. (1988). The person concept in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behaviour Processes*, 14, 43-55.
- Deputte, B. L., Pelletier, S., & Barbe, S. (2001). Visual categorization of natural and abstract items in forest monkeys and humans. *Behavioural Processes*, 55, 51-64. doi: http://dx.doi.org/10.1016/S0376-6357(01)00159-0
- Dunbar, R. I. M. (2003). The social brain: Mind, language and society in evolutionary perspective. *Annual Review of Anthropology*, 32, 163-181.
- Dungl, E., Schratter, D., & Huber, L. (2008). Discrimination of face-like patterns in the giant panda (*Ailuropoda melanoleuca*). *Journal of Comparative Psychology*, 122, 335-343. doi: 10.1037/0735-7036.122.4.335.
- Fagot, J., Martin-Malivel, J., & Dépy, D. (1999). What is the evidence for an equivalence between objects and pictures in birds and nonhuman primates? *Cahiers De Psychologie Cognitive/Current Psychology of Cognition*, 18, 923-949.
- Fize, D., Cauchoux, M., & Fabre-Thorpe, M. (2011). Humans and monkeys share visual representations. *Proceedings of the National Academy of Sciences, U.S.A.*, 108, 7635-7640. doi: 10.1073/pnas.1016213108.
- Fujita, K. (1987). Species recognition by five macaque monkeys. *Primates*, 28, 353 – 366.
- Gittleman, J. L. (1986). Carnivore brain size, behavioral ecology, and phylogeny. *Journal of Mammalogy*, 67, 23-36.
- Hare, B. (2007). From nonhuman to human mind: What changed and why? *Current Directions in Psychological Science*, 16, 60-64.
- Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The domestication of social cognition in dogs. *Science*, 298, 1634-1636.
- Hare, B. & Tomasello, M. (2005). Human-like social skills in dogs? *Trends in Cognitive Sciences*, 9, 439-444.
- Herrnstein, R. J. (1979). Acquisition, generalization, and discrimination reversal of a natural concept. *Journal of Experimental Psychology: Animal Behaviour Processes*, 5, 116-129.
- Herrnstein, R. J. (1990). Levels of stimulus control: A functional approach. *Cognition*, 37, 133-166. doi: 10.1016/0010-0277(90)90021-B
- Herrnstein, R. J., Loveland, D. H., & Cable, C. (1976). Natural concepts in pigeons. *Journal of Experimental Psychology: Animal Behaviour Processes*, 2, 285-302.
- Humphrey, N. K. (1976). The social function of intellect. In P. P. G. Bateson & R. A. Hinde (Eds.), *In growing points in ethology* (pp. 303-317). Cambridge, UK: Cambridge University Press.
- Jitsumori, M., & Matsuzawa, T. (1991). Picture perception in monkeys and pigeons: Transfer of rightside-up versus upside-down discrimination of photographic objects across conceptual categories. *Primates*, 32, 473-482.
- Jitsumori, M., Ohkita, M., & Ushitani, T. (2011). The learning of basic-level categories by pigeons: The prototype effect, attention, and effects of categorization. *Learning & Behavior*, 39, 271-287. doi: http://dx.doi.org/10.3758/s13420-011-0028-4
- Jitsumori, M., & Ohkubo, O. (1996). Orientation discrimination and categorization of photographs of natural objects by pigeons. *Behavioural Processes*, 38, 205-226.
- Jolly, A. (1966). Lemur social behaviour and primate intelligence. *Science*, 153, 501-506.
- Kelling, A. S., Synder, R. J., Marr, M. J., Bloomsmith, M. A., Gardner, W. & Maple, T. L. (2006). Color vision in the panda (*Ailuropoda melanoleuca*). *Learning & Behavior*, 34, 154-161.
- Kendrick, D. F., Wright, A. A., & Cook, R. G. (1990). On the role of memory in concept learning by pigeons. *The Psychological Record*, 40, 359-371.
- Kromrey, S., Maestri, M., Hauffen, K., Bart, E. & Hegdé, J. (2010). Fragment-based learning of visual object categories in non-human primates. *PLoS ONE*, 5, e15444. doi:10.1371/journal.pone.0015444

- Kubinyi, E., Virányi, Z., & Miklósi, A. (2007). Comparative social cognition: From wolf and dog to humans. *Comparative Cognition and Behavior Review*, 2, 26-46.
- Lazareva, O. F., Freiburger, K. L., & Wasserman, E. A. (2004). Pigeons concurrently categorize photographs at both basic and superordinate levels. *Psychonomic Bulletin & Review*, 11, 1111-1117.
- Lazareva, O. F., Soto, F. A., & Wasserman, E. A. (2010). Effect of between-category similarity on basic level superiority in pigeons. *Behavioural Processes*, 85, 236-245. doi:10.1016/j.beproc.2010.06.014
- Martin-Malivel, J. (1998). Existe-t-il une équivalence entre l'objet et sa représentation photographique chez le babouin (*Papio papio*). *Primateologie*, 1, 249-268.
- Martin-Malivel, J. & Fagot, J. (2001). Perception of pictorial-human faces by baboons: Effects of stimulus orientation on discrimination performance. *Animal Learning & Behavior*, 29, 10-20.
- Martin-Malivel, J., Mangini, M. C., Fagot, J., & Biederman, I. (2006). Do humans and baboons use the same information when categorizing human and baboon faces? *Psychological Science*, 17, 599-607. doi: http://dx.doi.org/10.1111/j.1467-9280.2006.01751.x
- Marsh, H. L., & MacDonald, S. E. (2008). The use of perceptual features in categorization by orangutans (*Pongo abelli*). *Animal Cognition*, 11, 569-585. doi:10.1007/s10071-008-0148-1
- Miklósi, A. & Topál, J. (2004). Review comparative social cognition: What can dogs teach us? *Animal Behaviour*, 67, 995-1004.
- Milton, K. (1981). Distribution patterns of tropical plant foods as an evolutionary stimulus to primate mental development. *American Anthropologist*, 83, 534-548.
- Milton, K. (1988). Foraging behaviour and the evolution of primate intelligence. In R. W. Byrne & A. Whiten (Eds.), *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans* (pp. 285-305). New York: Clarendon.
- Parron, C., Call, J., & Fagot, J. (2008). Behavioural responses to photographs by pictorially naïve baboons (*Papio anubis*), gorillas (*gorilla gorilla*) and chimpanzees (*Pan troglodytes*). *Behavioural Processes*, 78, 351-357.
- Perdue, B. M., Snyder, R. J., Pratte, J., Marr, M., & Maple, T. L. (2009). Spatial memory recall in the giant panda (*Ailuropoda melanoleuca*). *Journal of Comparative Psychology*, 123, 275-279. doi: 10.1037/a0016220.
- Perdue, B. M., Synder, R. J., Zhihe, Z., Marr, M. J., & Maple, T. (2011). Sex differences in spatial ability: A test of the range size hypothesis in the order Carnivora. *Biology Letters*, 7, 380-383.
- Premack, D. (1983). The codes of man and beasts. *Behavioral and Brain Sciences*, 6, 125-167.
- Range, F., Aust, U., Steurer, M., & Huber, L. (2008). Visual categorization of natural stimuli by domestic dogs (*Canis familiaris*). *Animal Cognition*, 11, 339-347.
- Rivas, J., & Burghardt, G. M. (2002). Crotalomorphism: A metaphor for understanding anthropomorphism by omission. In M. Bekoff, C. Allen & G. Burghardt (Eds.). *The Cognitive animal: Empirical and theoretical perspectives on animal cognition* (pp. 9-17). Cambridge, MA: MIT Press,
- Roberts, W. A., & Mazmanian, D. S. (1988). Concept learning at different levels of abstraction by pigeons, monkeys, and people. *Journal of Experimental Psychology: Animal Behaviour Processes*, 14, 247-260.
- Rosch, E. (1978). *Principles of categorization*. Cambridge, MA: MIT Press.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573-605.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382-439.
- Schrier, A. M., Angarella, R., & Povar, M. L. (1984). Studies of concept formation by stump-tailed monkeys: Concepts, humans, and letter A. *Journal of Experimental Psychology: Animal Behaviour Processes*, 10, 564-584.
- Schrier, A. M., & Brady, P. M. (1987). Categorization of natural stimuli by monkeys (*Macaca mulatta*): Effects of stimulus set size and modification of exemplars. *Journal of Experimental Psychology: Animal Behaviour Processes*, 13, 136-143.
- Shimp, C. P. (2004). Ambiguity, logic, simplicity, and dynamics: Wittgensteinian evaluative criteria in peer review of quantitative research on categorization. *Behavioural Processes*, 66, 333-348. doi: http://dx.doi.org/10.1016/j.beproc.2004.03.012
- Soto, F. A., & Wasserman, E. A. (2010). Error-driven learning in visual categorization and object recognition: A common-elements model. *Psychological Review*, 117, 349-381. doi: http://dx.doi.org/10.1037/a0018695
- Tanaka, M. (2001). Discrimination and categorization of photographs of natural objects by chimpanzees (*Pan troglodytes*). *Animal Cognition*, 4(3-4), 201-211. doi:10.1007/s100710100106
- Udell, M. A. R., Dorey, N. R., & Wynne, C. D. L. (2008). Wolves outperform dogs in following human social cues. *Animal Behaviour*, 76, 1767-1773.

- Udell, M. A. R., Dorey, N. R., & Wynne, C. D. L. (2010). What did domestication do to dogs? A new account of dogs' sensitivity to human actions. *Biological Reviews*, 85, 327-345.
- Udell, M. A. R., Dorey, N. R., & Wynne, C. D. L. (2012). Inter species social learning in dogs: The inextricable roles of phylogeny and ontogeny. In T. Zentall & E. Wasserman (Eds.), *Comparative Cognition*, 2nd ed. (pp. 819-831). New York: Oxford University Press.
- Vauclair, J., & Fagot, J. (1996). Categorization of alphanumeric characters by Guinea baboons: Within and between-class stimulus discrimination. *Current Psychology Cognition*, 15, 449-462.
- Vogels, R. (1999). Categorization of complex visual images by rhesus monkeys. part 1: Behavioural study. *European Journal of Neuroscience*, 11, 1223-1238. doi: <http://dx.doi.org/10.1046/j.1460-9568.1999.00530.x>
- Vonk, J. (2002). Can orangutans and gorillas acquire concepts for social relationships? *International Journal of Comparative Cognition*, 15, 257-277.
- Vonk, J. (2013). Matching based on biological categories in orangutans (*Pongo abelii*) and a gorilla (*Gorilla gorilla gorilla*). PeerJ. 1:e158 <http://dx.doi.org/10.7717/peerj.158>
- Vonk, J., & Beran, M. J. (2012). Bears "count" too: Quantity estimation and comparison in black bears (*Ursus americanus*). *Animal Behaviour*, 84, 231-238.
- Vonk, J., Jett, S. E., & Mosteller, K. W. (2012). Concept formation in American black bears (*Ursus americanus*). *Animal Behaviour*, 84, 953-964.
- Vonk, J., Jett, S. E., Mosteller, K. W., & Galvan, M. (2013). Natural category discrimination in chimpanzees (*Pan troglodytes*) at three levels of abstraction. *Learning and Behavior*, 41, 261-274. / 10.3758/s13420-013-0103-0
- Vonk, J. & Johnson – Ulrich, Z. (2014). Social and non-social category discriminations in a chimpanzee (*Pan troglodytes*) and American black bears (*Ursus americanus*). *Learning and Behavior*.
- Vonk, J., & MacDonald, S. E. (2002). Natural concept formation in a juvenile gorilla (*Gorilla gorilla gorilla*) at 3 levels of abstraction. *Journal of the Experimental Analysis of Behaviour*, 78, 315-332.
- Vonk, J., & MacDonald, S.E. (2004). Levels of abstraction in orangutan (*Pongo abelii*) categorization. *Journal of Comparative Psychology*, 118, 3-13.
- Vonk, J., & Mosteller, K.W. (2013) Perceptual versus conceptual memory processes in a chimpanzee (*Pan troglodytes*). In M. Howe, M. Toglia, H. Otgaar, H., & B. Schwartz (Eds.), *What is adaptive about adaptive memory* (pp. 258-267). Oxford University Press, New York.
- Vonk, J., & Povinelli, D. J. (2012). Similarity and difference in the conceptual systems of primates: The unobservability hypothesis. In E.Wasserman & T. Zentall (Eds.), *Oxford handbook of comparative cognition*, 2nd ed..(pp. 552-575) Oxford University Press, London.
- Vonk, J., & Shackelford, T. K. (2012). *Oxford handbook of comparative evolutionary psychology*. Oxford University Press., New York.
- Watanabe, S. (1993). Object-picture equivalence in the pigeon: An analysis with natural concept and pseudoconcept discriminations. *Behavioural Processes*, 30, 225-231.
- West, S., Jett, S. E., Beckman, T. & Vonk, J. (2010). The phylogenetic roots of cognitive dissonance. *Journal of Comparative Psychology*, 124, 425-432.
- Wynne, C. D. L., Udell, M. A. R., & Lord, K. A. (2008). Ontogeny's impact on human-dog communication. *Animal Behaviour*, 76.4, e1-e4.
- Yoshikubo, S. (1985). Species discrimination and concept formation by Rhesus monkeys (*Macaca mulatta*). *Primates*, 26, 285-299.
- Zamisch, V., & Vonk, J. (2012). Spatial memory in captive American black bears (*Ursus americanus*). *Journal of Comparative Psychology*, 126, 372-387.
- Zentall, T. R., Wasserman, E. A., Lazareva, O. F., Thompson, R. R., & Rattermann, M. (2008). Concept learning in animals. *Comparative Cognition & Behavior Review*, 313, 35-45. doi:10.3819/ccbr.2008.30002.