Title
Can Orangutans and Gorillas Acquire Concepts for Social Relationships?

Journal
International Journal of Comparative Psychology, 15(4)

ISSN
0889-3675

Author
Vonk, Jennifer

Publication Date
2002-12-31

DOI
10.46867/C4F30K

International Journal of Comparative Psychology, 2002, **15**, 257-277.

Can Orangutans and Gorillas Acquire Concepts for Social Relationships?

Jennifer Vonk
York University, Canada

Two orangutans' and one gorilla's understanding of social relationships was investigated using two-dimensional photographs displayed on a touch-screen monitor. Unlike the photos used in similar studies, the photos presented here were not of exclusively familiar or related individuals, thus eliminating the use of previously learned associations or "genetic" similarities as cues. In Experiment 1, the subjects discriminated photos of mother-offspring pairs from photographs depicting other social relationships (siblings, unrelated group mates, mated pairs). In Experiment 2, they matched photos of mother-offspring pairs, mated pairs, siblings and groups of animals in a delayed matching-to-sample task (DMTS). In Experiment 3, they matched photographs depicting various behaviors (sleeping, eating, playing and grooming) in another DMTS procedure. Performance was significantly above chance in all three experiments, suggesting that both species of Great Ape might be sensitive to abstract concepts such as social relationships and activities.

Typically, studies of social understanding in nonhuman primates rely upon observations of group dynamics and individuals' recognition of the dominance status of their conspecifics (Cheney & Seyfarth, 1999; Cheney, Seyfarth, & Smuts, 1986; Maestripieri, 1996). These studies have examined topics such as cooperative behaviors and individuals' abilities to understand the role of other individuals in problem-solving tasks (Call & Tomasello, 1995; Povinelli, Nelson, & Boysen, 1992; Povinelli, Parks, & Novak, 1992; Visalberghi, Quarantotii, & Tranchida, 2000), social foraging (Coussi-Korbel, 1994; Hare, Call, & Tomasello, 2001; Menzel, 1997), affiliative behaviour (Call, Laurel, & de Waal, 1999; Cheney, Seyfarth, & Silk, 1995; Silk, 1999), gaze following (Povinelli, 1999; Tomasello, Hare, & Agnetta, 1999) and social referencing (Russell, Bard, & Adamson, 1997). In the field, or in observational studies in general, it is difficult to rule out the role of learned associations between behaviors and outcomes or individuals. For instance, a subject may come to learn that the appropriate response to a dominant or aggressive individual is one of submission, without having any explicit representation of the concept of dominance. Submissive acts may become reinforced through the absence of aggression directed towards the observer by the dominant individual. Thus the observer comes to learn to behave appropriately in

This research was supported by a Natural Sciences and Engineering Research Council (NSERC) Postgraduate scholarship to the author. The cooperation and support of the staff at the Toronto Zoo was greatly appreciated. Special thanks to D. Devison, B. Carter, V. Phelan, G. Mintha, C. Guthrie, T. McCaskie, C. Wiederman, H. Minicke, M. Smith, K. Tunwell, B. Burke-Johnson, A. Beatson, R. Vos, D. Partington, D. Macguire, J. Armstrong, and J. Craig, without whose assistance this study would not have been possible. I am also grateful to S. MacDonald for the opportunity to undertake this project, and for providing the equipment, as well as for an insightful discussion regarding the writing of this article. Thanks also to B. Thompson and R. Sorge for helpful comments on an earlier draft of this paper. Correspondence concerning this article may be addressed to J. Vonk, University of Louisiana at Lafayette, 4401 West Admiral Doyle Drive, New Iberia, LA 70560, U.S.A. (jxv9592@louisiana.edu).

the presence of the dominant individual and, perhaps, generalizes this behavior to other individuals who appear or behave similarly to the dominant individual, without forming any kind of concept as to the nature of dominance.

A "true" concept is a general rule that can be applied to a variety of novel circumstances and does not depend upon learned associations between specific items or individuals (Premack, 1983). Whether nonhumans have a concept of social relationships is difficult to determine, particularly in an uncontrolled setting. A concept is considered to be abstract if it relies upon representations that are not tied to one or few physical features, and instead involves an understanding of how the category coheres based on many, sometimes not directly observable, features (Spalding & Ross, 2000). Furthermore, the use of a concept entails that the subject is able to transfer that knowledge to novel situations, or behave appropriately to stimuli with which he/she has had no prior experience. Because nonhumans do not have language, it is necessary to design nonverbal tests of their knowledge. Typically, this involves the presentation of pictorial stimuli. These stimuli necessarily contain various perceptual cues, and therefore make it challenging to draw conclusions about abstract or nonperceptually-based concepts. However, by testing for the presence of relational concepts, one requires that the subjects attend to the abstract relationship between the objects depicted and not solely to perceptual features.

Surprisingly few researchers have focused on the abilities of nonhumans to discriminate stimuli on the basis of the relationship between individuals represented in photographs. Boysen and Berntson (1989) demonstrated that a young female chimpanzee not only recognized individuals in slides, as evidenced by different heart-rate responses to familiar and unfamiliar conspecifics and caregivers, but also responded differently according to her relationship with the individuals portrayed in the slides. For instance, she showed accelerated heart-rate responses to slides of an animal that was aggressive towards her, while responses to slides of a playmate were minimal. That this chimpanzee's heart-rate responses corresponded to the nature of her social relationships with the individuals depicted in the slides suggested that chimpanzees represent photographs of conspecifics as depicting those familiar individuals, and that knowledge about the nature of their social interactions with those individuals can be evoked through the presentation of these two-dimensional stimuli.

Others have more explicitly tested the extent to which nonhuman primates form concepts about social relationships. Dasser (1988a, 1988b) demonstrated that Java monkeys could select photos of familiar mothers and their offspring from photos showing unrelated mother/young pairs of that species, and could also match photos of mothers to their offspring in a matching to sample (MTS) task. In Dasser's studies, subjects had to attend to information about the relationship between individuals in the photos, as well as to information about the identity of the subjects of the photographs. A more recent study (Parr & de Waal, 1999) provided evidence that chimpanzees could also match photos of unfamiliar mother chimpanzees to photos of their sons but not their daughters. This study was the first to show recognition of genetic relatedness between animals with which the subjects were unfamiliar, and thus eliminated the role of prior experience and familiarity. These studies show that nonhuman species can attend to identity and relatedness in their own species, in two-dimensional images. However, because related individuals presumably look more alike, and share more features in common than unrelated individuals, it is possible that these results were due to the pairing of stimuli that shared common physical features, as opposed to the formation of an abstract concept regarding the nature of the relationship depicted in the photographs.

The goal of the present research was to explore whether other Great Ape species might be capable of abstracting rules or concepts tying together stimuli on a nonperceptual basis. The current experiments may be the first to test the ability of nonhumans to extract the *general* nature of the social relationship depicted in photos of *unfamiliar* and sometimes *unrelated* animals of *various* species. Attending to physical characteristics of the stimuli in the current experiments would often hinder rather than facilitate performance. For instance, if the subjects attempted to learn associations between the stimuli based on cues such as color, relative size differences, number of individuals, and species of individuals in the photos, they would not perform the tasks accurately.

Photos depicting different social relationships, and which varied on several other dimensions, were presented to two male Sumatran orangutans *(Pongo abelii)* and one juvenile female western lowland gorilla *(Gorilla gorilla gorilla)*. One might expect gorillas to be more attuned to or captivated by social stimuli, given that they live in social groups, whereas orangutans are relatively solitary compared to other primate species. Thus if any species differences were to occur, one might expect the gorilla to more readily discriminate photographs on the basis of social relationships.

The subjects were first rewarded for selecting photos of mother and offspring pairs from photos of various other pairs of individuals, in a two-choice discrimination procedure. Transfer to photos of different species was then tested (Experiment 1). The apes were subsequently rewarded for MTS when shown photos exemplifying one of four possible social relationships, not always involving related individuals (Experiment 2). Finally, in Experiment 3, the ability of two of the subjects to attend to the behavior being performed by animals in photographs was tested in a delayed MTS task (DMTS).

Experiment 1

It was difficult to select photos that instantiated an obvious social relationship. The relationship between mother/offspring pairs represents a clear case of an important and adaptive social relationship. Thus, sets of photos of mother/offspring pairs were chosen that could be contrasted with photos of adults and infants, either as individuals, in pairs, or in groups. The task was a two-choice discrimination procedure in which the subjects were presented with pairs of photos and required to select the "correct" photo (in this case photos of mother/offspring pairs) for reinforcement. The first and second photo set (training and first transfer photos) included photos of only one species, orangutans. The third photo set comprised a transfer test to a different species, gorillas. The fourth and fifth transfer sets consisted of photos of a mixture of primate species, and the sixth and final transfer set was comprised of photos of various nonprimate animals. If the subjects understood the concept being tested, that is the mother/offspring relationship, they should continue to show positive transfer to each new photo set despite the change in species depicted in the photos.

Method

Subjects. Two adult male Sumatran orangutans, Molek, age 22, and Dinar, age 13, and one juvenile female western lowland gorilla, Zuri, age 4, served as subjects in this experiment. All were housed at the Toronto Zoo, Toronto, Ontario, Canada. Zuri was housed individually and Molek and Dinar were housed individually when they were not on exhibit with four other orangutans during the day, two or three days a week. Zuri was human reared and arrived at the Toronto Zoo at the age of 18 months. Molek had been previously housed at Yerkes Primate Center. Dinar was reared by his mother at the Toronto Zoo.

All subjects had previously participated in experiments using the touch screen (Vonk, 2003a, 2003b in press; Vonk & MacDonald, 2002, in press). They had been trained to choose (by touching) one photo when pairs of photographs were presented, and were rewarded for selecting only the photo belonging to the reinforced category. Discriminations involved choosing between photos of orangutans and humans, orangutans and other primates, primates and other animals and animals, versus nonanimals (Vonk & MacDonald, 2002, in press). The gorilla had also learned to discriminate between photos of gorillas and humans and gorillas and other primates (Vonk & MacDonald, 2002). On initial presentation of the touch screen they were rewarded for selecting any photo by touching anywhere within its boundaries. Once subjects were reliably selecting the photos (typically after the first two sessions), they were rewarded only for selecting the photos belonging to the reinforced categories. On average, fewer than 20 sessions were required in order for the subjects to achieve criterion of two successive sessions at 80% correct.

Each subject had also participated in DMTS tasks in which they learned to match photos according to species (Japanese macaques, gibbons, lemurs, golden lion tamarins, or proboscis monkeys) and taxonomic categories (birds, reptiles, insects, mammals, or fish; Vonk, 2003b). They were rewarded for touching the sample photo and its correct match for the first two sessions and subsequently were rewarded only for correctly selecting the comparison photo that matched the sample, even on novel DMTS tasks with novel stimuli. They had seen various photographs of some of the same species used in the present experiment, but each photo was itself novel. Subjects had never participated in identity MTS tasks.

Materials. Six sets of 20 color photos each were created (120 photos in total). Each set contained completely novel photos. As stated above, the first and second photo set (training and first transfer photos) included photos of only one species, orangutans. Some of the orangutan photos were photos of individuals familiar to the orangutan subjects. Half (10) of the photos in both sets were of a mother and infant orangutan. These were the positive or reinforced exemplars. The other ten photos (the negative or nonreinforced exemplars) consisted of a mixture of individuals, pairs, or groups of orangutans of various genders and ages in various postures and with differing backgrounds. Backgrounds were matched as much as possible between the positive and negative exemplars. The negative exemplars included photos in which pairs or groups of animals differed in size and age, thus making it difficult to rely on this perceptual pattern to make the discrimination. Examples of positive and negative stimuli are depicted in Figure 1.

The third photo set comprised a transfer test to a different species, and included only photos of gorillas, making successful transfer slightly more challenging. Again, half of the photos were of mothers with infants and the other, nonreinforced, photos were of either individuals or pairs of gorillas of various ages and genders. One photo showed an adult male silverback with an infant and should have been difficult to discriminate perceptually from the mother/offspring photos, especially for the orangutan subjects who had never seen live gorillas.

The fourth and fifth transfer sets consisted of photos of a mixture of various primate species. The negative exemplars in the fifth set were always pairs of subjects, as opposed to including some individuals, thus making it a more difficult set than the fourth set. Both sets included some photos of the same species between the negative and positive exemplar groupings and therefore the subject would sometimes choose incorrectly if he/she mistakenly attended to species versus social information. The subjects had previously been trained to attend to only taxonomic information (Vonk, 2003b; Vonk & MacDonald, 2002, in press) so such errors would not be unexpected.

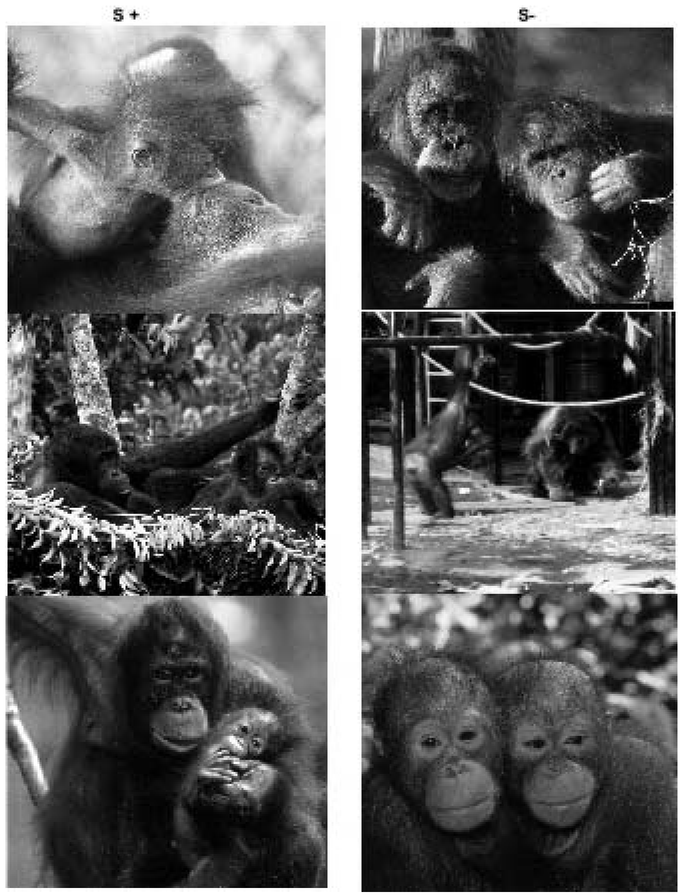

Figure 1. Examples of both S+ and S- stimuli from the first set of mother/offspring photographs in Experiment One. Positive exemplars appear on the left and negative exemplars appear on the right.

The sixth and final transfer set was comprised of photos of various nonprimate animals from several taxa. This set included some species never before seen by any of the subjects. All nonreinforced photo sets included some photos of young animals and some photos of adult females so that the subjects could not perform accurately simply by choosing photographs of juveniles and/or adult females.

Procedure. The experiment was programmed in Filemaker Pro 3 software and was run on a Macintosh 5300 PowerBook computer. The digitally scanned photos were presented on a 13 inch Apple touch screen monitor, and were approximately 3 x 4 inches on the screen. The monitor was placed against the bars of the subjects' cages and the subjects were required to either reach through the mesh holes to touch the screen (Molek and Dinar) or to reach underneath and touch the monitor (Zuri). The experimenter was positioned to one side of the screen, which was completely covered in a wooden protective covering. Thus she could not see the stimuli projected on the screen and viewed the experiment on the PowerBook placed behind the monitor in order to determine when to present reward. Animals could see the experimenter during trials.

Subjects were tested individually at the same time each day. Molek and Dinar received one to four sessions per day three or four times a week while Zuri was given between five and twelve sessions a day four days a week. This discrepancy was due to the availability of the subjects and keepers. However, no more than four of these sessions involved the same photographs on the same test day. Sessions of this experiment were sometimes interspersed with sessions from other unrelated experiments. For instance the subjects also participated in tests of their understanding of the relationships between components of abstract stimuli in which they had to match images according to whether two components of a stimulus shared the same shape or same color (Vonk, 2003a). In addition, they were tested in a variety of two-choice discriminations similar to that described in Vonk and MacDonald (2002). Molek and Dinar also participated in DMTS tasks designed to determine if they could match images based on the number of dots or individuals depicted in the stimuli (unpublished data).

Each session consisted of presentation of 10 pairs of photos (10 trials, 20 photos). Within each pairing, which was randomly determined on each trial, was a photo of a reinforced and a nonreinforced exemplar. Side of presentation was counterbalanced so that half the positive and half the negative stimuli appeared on each side. The order of presentation of the photographs was randomized on each session. Each photo appeared only once within each session. The same photos appeared on each session with the same photo set. The subject was rewarded for selecting (by touching) photos of mother/offspring pairs or not rewarded if he/she made an incorrect choice. Touching one of the photos caused the program to advance to a screen that informed the experimenter of the subject's response. There was no time out for an incorrect response. Reinforcement consisted of a small piece of a highly preferred food (M&M, a nut, or dried fruit). The session continued until all 10 trials had been completed. When eight or more correct choices were made on two consecutive sessions, presentation of the next set of novel photographs commenced, typically on the same day but occasionally on the following test day. Performance on only the first two sessions with novel photographs was taken as an indication of transfer performance, but presentation of those photos continued until criterion was reached again.

Because of the small number of subjects and the differences in their ages, it was not appropriate to compare performance across the species tested in these experiments. Therefore separate analyses were conducted for each subject. In all the statistical tests reported in this article, alpha was set at the 0.05 level.

Results and Discussion

Chi-square tests comparing the total percentage correct for the last two sessions of each of the six photo sets (criterion sessions) to the first two sessions with the five sets of novel photos (transfer sessions), were conducted, separately for each subject, to determine whether transfer performance differed from criterion. These tests revealed no significant difference for Molek, $\chi^2(4) = 2.80$, but a significant difference for Dinar, $\chi^2(3) = 9.55$, and for Zuri, $\chi^2(3) = 7.82$, suggesting that transfer was equivalent to criterion performance only for Molek. Performance was lower on transfer relative to criterion trials for Zuri and for Dinar.

One-sample t-tests determined that transfer performance (the first two sessions with each of the five novel sets taken together) was significantly above chance (50%), for each subject, [Molek: $t(9) = 18.58$, Dinar: $t(9) = 22.42$, and Zuri: $t(9) = 24.71$]. Therefore, although Dinar's and Zuri's transfer performance was not equivalent to that reached at criterion, it was significantly above chance for both subjects.

It was possible that subjects could have responded correctly by attending to factors such as size differences between the individuals in the photos. Various factors were identified that might enable the subjects to make the discrimination. These relevant factors were size (relative size difference between the individuals in the photos, when there was more than one individual), number (number of individuals in the photo), presence of infant, presence of an adult female, and relative age of the offspring (infant, young juvenile, or slightly older juvenile). A series of Chi-square tests comparing the values that the coders assigned to each photo for negative and positive exemplar sets determined that the positive and negative exemplar sets could be differentiated on the basis of the following cues: size [$\chi^2(2, 95) = 49.64$], age [$\chi^2(2, 90) = 11.40$], the presence of a young animal,

[χ^2(1, 120) = 29.87], the presence of an adult female [χ^2(1, 120) = 60.00], the number of individuals in the photo [χ^2(5, 120) = 26.98], and whether or not there was more than one species of animal depicted [χ^2(1, 120) = 5.22]. Thus, it was possible that accurate performance on the task was achieved by attending to one or several of the aforementioned cues.

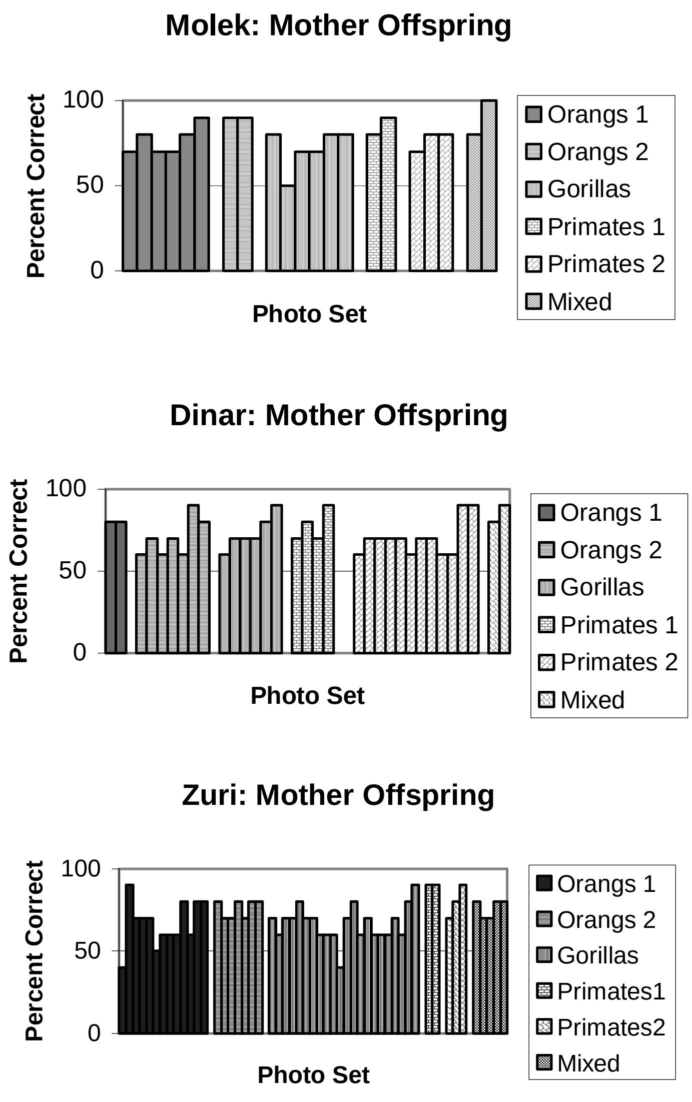

Figure 2. Percentage correct across sessions for each set of photos for Molek (top), Dinar (middle) and Zuri (bottom) in Experiment One. Each bar represents the percent correct for a single session comprised of 10 trials.

Hence, univariate analyses of variance were conducted to determine which of the factors might have controlled responding in the task. The scores for all factors (even those whose values did not differ between negative and positive photo sets) were included in the analysis as independent variables (main effects only). The percentage of correct responding to individual photos was transformed to log scores for each subject and these transformed scores served as the dependent variable. For positive exemplars, correct performance involved selecting the photo; for negative exemplars, correct responding involved *not* selecting that photo. Dinar and Molek were both significantly affected only by the presence of an adult female in the images, $F\text{s}(1, 40) = 33.16$ and 4.29 for Dinar and Molek, respectively. Both were more accurate when there was an adult female in the image. Zuri's performance was affected by the size difference between the individual in the photos, $F(2, 40) = 4.00$, and the proportion of the individual shown in the photos, $F(9, 40) = 2.49$. She was most accurate when there was a large size difference between the individuals depicted and when the images of the individuals were relatively large. Thus it is possible that cues other than the general concept of "mother" dictated the subjects' accuracy in this task.

All three subjects met criterion in less than 15 sessions (range 2 to 13) and showed a high degree of transfer to novel stimuli. Molek required fewer sessions to reach criterion and showed transfer that did not differ from criterion performance, while Zuri's and Dinar's transfer was significantly above chance. It seems unlikely that the use of single factors to discriminate the stimulus sets would lead to such high degrees of transfer as were observed here. However, it was not possible to rule out a strategy of attending to various combinations of features in the stimuli.

Zuri required the greatest number of sessions to reach criterion with the third set of photos. This was surprising because this set was comprised of photos of her own species, gorillas. Analysis of her errors revealed that she most often mistakenly selected photos of other juvenile gorillas, even though these responses were never reinforced. It appeared that she simply preferred photos of other solitary young individuals of her own species. However, when shown photos of other species she was able to select photos of young animals only in the context of the mother/offspring relationship and to avoid selecting solitary young animals. Across all photo sets, the presence of a young animal was not a significant factor in her responding.

All three subjects were able to transfer high levels of performance to novel photos of unfamiliar species. Because they had been trained to make discriminations based on distinctions between members of different species in previous studies, the finding that performance remained high in a task in which species was irrelevant is particularly impressive. The results indicate that these subjects were able to rapidly apply a new rule in this experiment and were able to attend to a concept less perceptually bound (i.e., one that did not rely purely on single or few perceptual features). Although it can be argued that they may have been using a rule or series of rules such as "pick photo of big and small animal of same species" (some photos included two individuals of different species, such as an adult female orangutan with a young human child), the application of rules is not necessarily incompatible with a concept interpretation. Indeed, according to feature based theories of concept formation, it is presumably through weighing the presence and absence of various features that humans are able to assign novel stimuli to categories (Hampton, 1981; Tversky, 1977). Detailed analysis of the features associated with each photo and pattern of responding to those photos by each subject revealed that very few of the single features that might have lead to correct performance on this task controlled the subjects' responding. Two of the subjects were affected only by the presence of an adult female in the photographs.

Because adult females appeared in some of the negative exemplars as well as in all of the positive exemplars it seems unlikely that relying on this cue alone would lead to transfer of 80% or more on novel sets, as was observed. Zuri apparently attended to the size difference between the two individuals. This was another relevant perceptual cue that might have contributed to a high level of accuracy in the task. While it remains possible that reliance on this cue alone lead to Zuri's high levels of performance, it also seems plausible that subjects might have been using several cues in conjunction in order to achieve such high levels of performance.

Perhaps the subjects did not learn the "mother/offspring" concept in general, but instead learned to use a rule such as "pick photo which includes image of adult female and smaller infant of same species." It is possible that concepts are formed through this very process of tying various rules together into a coherent and stable category representation (Medin & Smith, 1981). Thus, it is possible that subjects learned concepts through the course of the experiment and did not map the photos on to preexisting concepts for mother/offspring relationships.

Interestingly, the orangutans and the gorilla often differed in terms of which individual photos they selected. Most significantly, the two orangutan subjects consistently incorrectly chose the photo of a silverback gorilla with a very young gorilla (83.3 and 100 % of the time for Molek and Dinar, respectively), while Zuri correctly avoided selecting this photo 86.4 percent of the time. For subjects who did not recognize the silverback as a male, this photo would appear to depict a mother and offspring, thus difficulty with this photo was expected. If the orangutans understood the concept being tested but did not recognize the silverback as a male they should incorrectly select this photo often, as they did. Zuri however seemed able to discriminate between this photo and the mother/offspring photos. This finding would support the idea that she both understood the nature of the concept being tested and recognized the silverback as a male. Of these subjects, only Zuri had experience with live male and female adult gorillas.

Experiment 2

Next, the subjects were presented with a delayed matching-to-sample (DMTS) procedure in which the strategy of choosing photos on the basis of single or few perceptual features would be made even less efficient. Thus, lower levels of performance in this experiment might suggest such a strategy. Moreover, the use of a DMTS procedure made it possible to test for four different social concepts within a single session.

A sample stimulus was presented until the subjects attended to it and touched it, and then it disappeared. After a brief delay, two comparison stimuli appeared on the screen. Four different types of social relationship photos served as stimuli. One type was novel instances of mother/offspring pairs. The other types of discrimination were "social groups" of animals, mated pairs, and siblings. The latter two types involved pairs of individuals. In order to perform accurately on these trials, the subjects had to attend to cues such as the age of the individuals depicted and sexual dimorphism, or perhaps use the behavior of the animals depicted in the photos as clues. For instance, if the photo depicted two individuals mating, it might be more easily classified as a mated pair than if the photo depicted two individuals playing. The latter instance might be more readily classified as a portrayal of siblings. The test discouraged the use of perceptual features, in light of the fact that sometimes the species depicted in the nonreinforced comparison matched that depicted in the sample, while the species depicted in the reinforced comparison stimulus (the one instantiating the same social relationship as that in the sample) would differ from that depicted in the sample. For example, the sample might depict a mated chimpanzee pair and the comparison stimuli might include an image of chimpanzee siblings and an image of a mated pair of birds. In this example, the chimpanzee siblings would appear more similar to the sample image of chimpanzee mates, however this choice would be incorrect. The subjects would have to disregard information about the species or appearance of the individuals in the sample photo and attend to information about the type of relationship between those individuals. The correct response in this example would be to select the photo of the bird pair. Out of 240 total trials, the species in the correct comparison matched the species in the sample on only 25 trials, while the species in the incorrect comparison matched the species in the sample on 44 trials.

Method

Subjects. The same three subjects participated in this experiment.

Materials. 24 novel color photographs were presented. There were six photographs within each category (mother/offspring, group, mated pair, and siblings). The photos within each category were a mixture of animals from various taxonomic categories, such as birds, reptiles, and primates. Some species appeared in more than one category. The same 24 photographs were used on each session.

Different aspects of the photos were controlled so that the use of strategies other than conceptual categorization could not lead to highly successful performance on this task. For instance, a difference in size between the individuals might be expected for the mother/offspring photos. However, some photos from each of the other categories also depicted individuals of varying sizes, so that this variable alone could not distinguish these photos from those belonging to the other categories. In addition, siblings were always the same color as each other, but this was also true of most of the photos from each of the other categories, so that a strategy of choosing two individuals of the same (or different) color could not lead to accurate performance. It might also be expected that only siblings would display “play faces,” yet only three of the six sibling photos showed animals displaying “play faces.” Furthermore, similar expressions also appeared within the “mated pair” and “social group” photo categories. Therefore facial expression was unlikely to be a useful cue. Amount of contact between individuals in the photos also varied within the categories so could not be used as a reliable cue. One cue that was not controlled was the number of individuals in the photos. All of the categories except for the “social group” category included photos that showed only two individuals. Thus number is a concept that could also be used by the subjects to accurately match the “social group” photos, without necessarily making use of the social concepts being tested.

Procedure. The experiment was conducted in the same location and in the same manner as Experiment 1. Each session consisted of twelve trials. Therefore each block of five sessions consisted of 60 trials. During a trial, a sample photograph was presented in the center of the screen and stayed on screen until the subject attended to the photo and touched it. The two comparison photos subsequently appeared on the screen after a brief delay (approximately 3 s). The subject was then required to select, by touching, only one of the two comparison photos. If he or she selected the photo that matched the social relationship depicted in the sample he or she was given a small food reward (M&Ms or dried fruits and nuts) by hand. After selection of the comparison photo and reward, if required, the screen advanced to the next sample photo until all twelve trials were completed. There was no time out for incorrect choices.

During the 12 trials within a session, each of the four types of relationships was depicted in the sample photo 3 times, with a different exemplar from the category being used for each of the 3 trials. Each category was tested against each of the other 3 categories during each session, so that there would be, for instance, 3 trials where siblings appeared in the sample photo. Different sibling photos were used as the correct matches (or reinforced comparison stimuli) for these samples on all three of these trials. There were no identity matching trials. On one "sibling" trial, the nonreinforced comparison stimulus depicted a mated pair. On another "sibling" trial, the nonreinforced comparison stimulus depicted a group of animals and, on the third trial, the nonreinforced comparison stimulus depicted a mother/offspring pair, and so on for each discrimination type.

Three photographs within each category appeared twice within each session, while the other three photographs within each category appeared only once per session. The photos were sorted randomly before each session. Therefore, the photos that appeared twice on some sessions appeared only once on other sessions. If a photo appeared twice within a session, it appeared as both a reinforced and a nonreinforced comparison stimulus on different trials, but not as a sample. Therefore a photo that was correct once would be incorrect on its second presentation or vice versa. The photos that appeared as samples during a session appeared only once during that session. Side of presentation was counterbalanced so that half the correct choices appeared on the left and half appeared on the right side of the screen. Order of presentation and pairing of stimuli was random. Therefore each session involved many novel pairings of stimuli within trials.

Zuri and Molek received six blocks of five sessions for a total of 30 sessions (360 trials) and Dinar received four blocks of five sessions for a total of 20 sessions (240 trials).

Results and Discussion

The primary data appear in Figure 3. Binomial tests comparing performance across all sessions to chance (50%) revealed significant above-chance responding by each subject, for Dinar: N = 240, for Molek and Zuri: Ns = 360. Binomial tests were also conducted to determine when acquisition of above chance levels of responding occurred. Both Molek and Zuri were above chance by only the second session, Ns = 24, while Dinar's performance was above chance by the third session, N = 36. These results indicated that individual stimulus/response associations were not learned over prolonged training.

In order to assess whether substantial learning occurred throughout the course of the experiment, paired *t*–tests were conducted to compare performance on the first to the last block of testing. Performance did not improve significantly from the first block to the last block of sessions for any of the subjects, highest $t(4) = -1.81$.

In order to rule out the possibility that high levels of performance were simply due to rapid learning of associations between specific exemplars, performance on novel pairings was compared to performance on previously paired stimuli for the first six sessions only. Paired t-tests indicated that performance did not differ between novel and previous pairings for any of the subjects, highest $t(4) = 0.78$.

One sample *t*-tests were also conducted to determine whether performance on both novel and previous pairings was above chance (50%), also for the first six sessions only. For Dinar, performance on both novel and previous pairings was significantly above chance, $t(5) = 2.73$ and $t(4) = 3.20$ respectively. For Molek, only performance on previous pairings was above chance, $t(4) = 7.26$. For Zuri, performance on neither novel nor previous pairings was above chance, highest $t(4) = 1.83$.

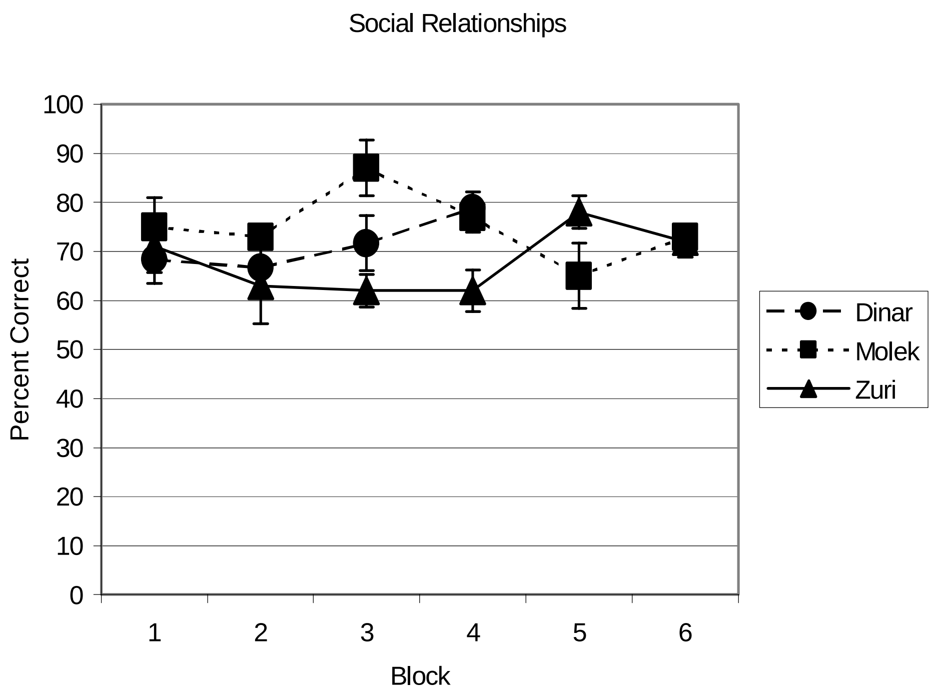

Figure 3. Average percentage correct across blocks of 5 sessions in Experiment 2.

It was of interest to determine whether some types of discriminations were made more readily than others. Chi-square tests comparing the percentage correct according to the different types of relation tested revealed significant effects for both Molek, $\chi^2(3) = 10.99$, and Zuri, $\chi^2(3) = 47.16$. However, the ease with which each subject performed the different types of discrimination differed. The type of relation depicted in the sample photo did not significantly affect Dinar's performance, $\Pi^2(3) = 3.03$. The data can be seen in Table 1.

Table 1
Subjects' Performance Across Different Types of Relation Tested in Experiment 2.

Subject	*Type*	*Mean*	*SD*	*Binomial p*
Dinar	Mother/Offspring	.72	.43	< .001
	Siblings	.76	.45	< .001
	Mated Pair	.63	.37	.05
	Social Group	.75	.46	< .001
Molek	Mother/Offspring	.83	.37	< .001
	Siblings	.79	.41	< .001
	Mated Pair	.75	.44	< .001
	Social Group	.63	.49	.01
Zuri	Mother/Offspring	.92	.27	< .001
	Siblings	.76	.43	< .001
	Mated Pair	.50	.50	.54
	Social Group	.54	.50	.23

Social group photos might have been discriminated using a process such as enumerating the number of individuals in the photos, as this relation was the only one in which more than two individuals appeared (Molek has since been tested on "match to the number of individuals" present in a photograph and performed above chance from the beginning (unpublished data), suggesting that he is indeed sensitive to numerosity). Yet these photos were matched poorly. If the subjects were looking for cues relevant to social relationships, this result is not surprising because the relationship in these photos was somewhat more ambiguous. Other aspects of the photos were controlled so that subjects could not use size or color differences between individuals, specific facial expressions, or amount of contact between individuals as cues to guide performance. Therefore the results are consistent with the idea that the subjects may have used social concepts to guide performance, as opposed to using more perceptually based strategies.

However, Zuri's results suggest that she may have simply preferred photos of young animals and therefore often selected photos of mother/offspring pairs and siblings regardless of whether they were appropriate matches to the sample or not. This bias would explain her higher performance on these discriminations relative to the others.

Although the subjects had previously been trained to make discriminations based on species categories (Vonk, 2003b; Vonk & MacDonald, 2002, in press), they did not appear to be matching photos on the basis of the species appearing in the sample and comparison photos. For each trial it was recorded whether or not each of the correct and incorrect comparison stimuli matched the species of the sample. As can be seen in Table 2 the subjects did not tend to be more accurate when the correct stimuli matched the species of the sample or when the incorrect stimuli did not match the species in the sample. This result argues against the idea that subjects were likely to match photos on the basis of physical features, as animals of the same species would presumably appear more alike than animals of a different species, and yet same species photos were not always matched at greater frequencies.

Table 2
Subjects' Performance According to Whether the Same Species was Depicted in Both the Sample and Each of the Two Comparison Stimuli in Experiment 2.

Subject	*Match*	*Species*	*Mean*	*SD*	*Binomial p*
Dinar	Sample to Correct Stimulus	Same	.68	.48	< .001
		Different	.72	.45	< .001
	Sample to Incorrect Stimulus	Same	.70	.46	< .001
		Different	.71	.45	< .001
Molek	Sample to Correct Stimulus	Same	.76	.43	< .001
		Different	.75	.43	< .001
	Sample to Incorrect Stimulus	Same	.62	.49	.01
		Different	.78	.41	< .001
Zuri	Sample to Correct Stimulus	Same	.58	.50	.07
		Different	.70	.46	< .001
	Sample to Incorrect Stimulus	Same	.62	.49	.01
		Different	.69	.46	< .001

The orangutans' overall performance was consistently well above chance for all types of discriminations within the first three sessions, and across all blocks, suggesting that they did not have to learn to associate various exemplars with each other in order to match accurately. Also in support of this conclusion, performance did not differ according to whether or not specific exemplars had been paired on previous trials or comprised novel pairings.

Although the use of several perceptual cues would necessarily aid performance, it is unlikely that any single feature could have contributed enough information to allow for the high levels of performance observed here. The subjects would need to make use of several cues to gain an understanding of the categories to be discriminated. The combining of features into a coherent category is what defines a concept as something more than an observable perceptual attribute (Spalding & Ross, 2000).

The results again suggest that these adult male orangutans may have attended more to social information than this young female gorilla did. However, the gorilla subject was much younger than the orangutans and lacked social experience with her peers as well. Therefore it would not be appropriate to speculate about species differences at this point.

Experiment 3

The subjects, in particular the orangutans, performed at consistently high levels in the previous experiments, suggesting that they might have been sensitive to information about social relationships presented pictorially. Behavior is one is one cue that might be used to understand the nature of social interactions. In order to categorize photographs according to what an individual is doing, versus what an individual looks like, an individual would have to represent functional categories. Such categories might be of relevance in social situations. This experiment examined the ability of two of the subjects to match photos based on the behavior being executed by the subjects of the photos. Subjects were asked to match-to-sample photos of sleeping, grooming, eating, or playing animals of various species in a DMTS task. To my knowledge, no other experiment has tested this discrimination in any species.

Method

Subjects. Molek and Zuri participated in this experiment.

Materials. 24 novel color photographs were presented. There were six photographs within each of four categories (eating, sleeping, grooming, and playing). The photos within each category represented a mixture of animals from various taxonomic categories, but were predominantly of primates. Several species appeared in more than one category.

Certain features of the stimuli were controlled so as to reduce the use of perceptual cues to guide performance in this task. For instance, food was sometimes present in photos where animals were not explicitly feeding so that “eating” could not be detected simply by the presence of food. In addition, photos of animals with their eyes closed were included in the “playing” and “grooming” categories so that closed eyes were not a reliable cue for detecting “sleeping” photos. Expressions like the “play face” appeared in both “playing” and “grooming” photos so could not be used to discriminate between these two categories. Furthermore, not all “playing” photos showed animals displaying that expression. Posture, amount of contact between individuals, and number of individuals also varied within and across the different categories.

All 24 photographs appeared during each session and the same photographs were used in each session.

Procedure. The procedure was identical to that used in Experiment 2, except for the nature of the discriminations tested and the materials described above. Both subjects received four blocks of five sessions for a total of 20 sessions (240 trials).

Results and Discussion

As shown in Figure 4, and confirmed by binomial tests, performance was above chance (50%) across all sessions, for both Molek and Zuri, *N*s = 240. Binomial tests also revealed that above chance levels of responding were acquired by the first session for Zuri, *N* = 12, and by the third session for Molek, *N* = 36. These results again indicated that the subjects did not require prolonged exposure to the stimuli to learn the discriminations.

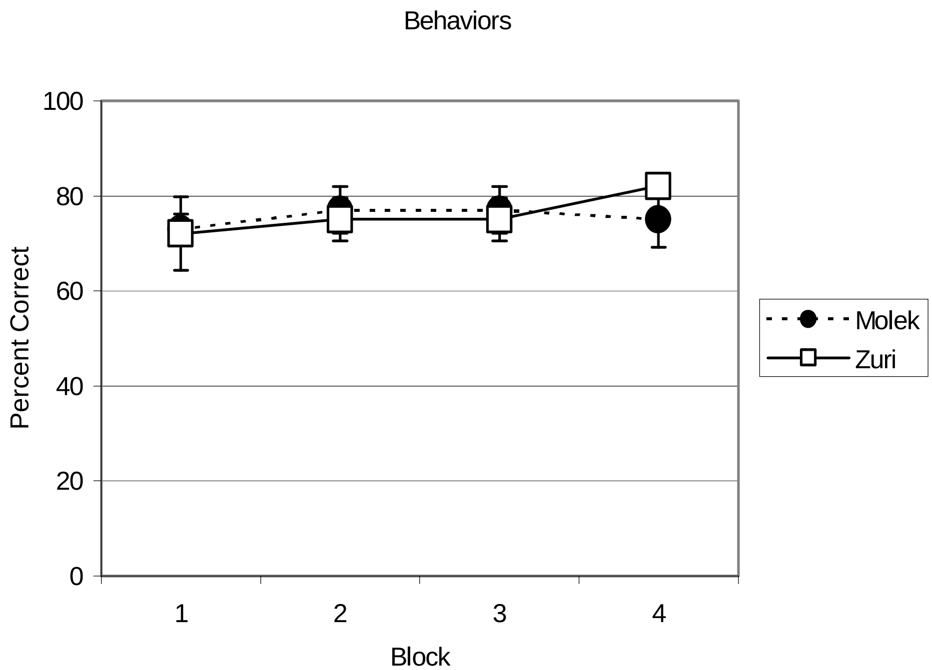

Figure 4. Average percentage correct across blocks of 5 sessions in Experiment 3.

In order to demonstrate that substantial learning did not occur throughout the course of the experiment, paired *t*–tests were conducted to compare performance on the first to the last block of testing. Performance did not improve from the first block to the last block of sessions for either of the subjects, highest *t*(4) = -1.51.

In order to rule out the possibility that high levels of performance were simply due to rapid learning of associations between specific exemplars, performance on novel pairings was compared to performance on previously paired stimuli for the first six sessions only. Paired *t*-tests indicated that performance differed between novel and previous pairings for Molek alone, *t*(4) = 4.78. His performance was more accurate with previous pairings.

One sample *t*-tests were also conducted to determine whether performance on both novel and previous pairings was above chance (50%), also for the first six sessions only. For Molek, performance on both novel and previous pairings was significantly above chance, *t*(5) = 7.19 and *t*(4) = 13.06. For Zuri, only performance on new pairings was above chance, $t(5) = 5.03$. For Molek, performance was superior on pairs that included repeated pairings, yet for both subjects, performance on novel pairings was significantly above chance, suggesting that learning associations between previously paired stimuli was not necessary for accurate performance.

It was possible that certain behaviors were recognized and matched with greater accuracy compared to other behaviors. Chi-square tests comparing each subject's percentage of correct responses according to the discrimination type (the behavior shown in the sample photo) revealed that Molek's performance varied as a result of type of discrimination, $\chi^2(3) = 21.82$, while Zuri's did not, $\chi^2(3) = 4.09$. The data appear in Table 3.

Table 3
Subjects' Performance Across Different Types of Behavior Tested in Experiment 3.

Subject	*Type*	*Mean*	*SD*	*Binomial p*
Molek	Sleeping	.90	.30	< .001
	Eating	.87	.34	< .001
	Grooming	.62	.49	.05
	Playing	.63	.49	.03
Zuri	Sleeping	.85	.36	< .001
	Eating	.70	.46	.001
	Grooming	.75	.44	< .001
	Playing	.73	.45	< .001

Both subjects discriminated the photos on the basis of the behaviours depicted in the photographs from the first session onward, and showed a fairly high degree of consistency across sessions and discriminations, although Zuri was more consistent than Molek with regard to the type of behaviours depicted. The results suggest that both species attend to more than perceptual information in photographs and do so without extensive training. Again, information about the species of the animals appearing in the photos was irrelevant to the task and attending to such information would have led to errors. Such errors would have led to lower levels of performance than were observed here.

Once again, information about whether or not the correct and incorrect comparison stimuli matched the species in the sample was recorded and analyzed. There were 135 trials on which the correct comparison matched the species in the sample, and 90 trials on which this was true for the incorrect comparison. For Molek, this time, whether or not the species depicted in the sample was the same as the species depicted in the correct comparison, affected his performance. He was more accurate when the species did match. However, he also tended to be more accurate when the incorrect sample matched the species in the sample, indicating that he was not strictly biased to select the stimulus that matched the sample in terms of the species depicted. Whether or not the species matched between the sample and either of the comparison stimuli did not influence Zuri's performance. Surprisingly, she tended to be less accurate when the correct comparison depicted the same species as the sample. Again these results suggest that neither subject merely tended to match photos on the basis of which comparison looked more like the sample. The data appear in Table 4.

Table 4
Subjects' Performance According to Whether the Same Species was Depicted in Both the Sample and Each of the Two Comparison Stimuli in Experiment 3.

Subject	*Match*	*Species*	*Mean*	*SD*	*Binomial p*
Molek	Sample to Correct Stimulus	Same	.91	.24	< .001
		Different	.70	.46	< .001
	Sample to Incorrect Stimulus	Same	.83	.38	< .001
		Different	.73	.44	< .001
Zuri	Sample to Correct Stimulus	Same	.70	.46	< .001
		Different	.78	.42	< .001
	Sample to Incorrect Stimulus	Same	.80	.41	< .001
		Different	.75	.43	< .001

It is possible that attending to information about the behaviors involved in the photos aided in the previous task, where the subjects were required to distinguish between siblings, who were often shown at play, and mated pairs, who might be shown mating or grooming in some, but not all, instances. In this experiment, the gorilla's performance was more consistent and less influenced by irrelevant features, compared to the orangutan's performance.

General Discussion

Clearly both the orangutans and the gorilla in these experiments were able to make the discriminations more accurately than would be predicted by chance. They rapidly reached high levels of performance and transferred to new photos at levels well above chance in Experiment 1. Attending to obvious irrelevant features, such as the number of individuals and species of animals in the photos, would lead to lower levels of performance. It is of course possible that other cues not necessarily apparent to the researcher, or combinations of various features, may have been guiding performance in these tasks. However, because photos were controlled for various factors such as facial expression, amount of contact, and size differences, it seems unlikely that attending to other cues would have led to such high levels of accuracy.

Prior research has shown that other species are capable of extracting information about social roles in group situations. Some have shown that nonhuman primates can recognize unfamiliar individuals in photographs (Parr, Winslow, Hopkins, & DeWaal, 2000), and match these photos to photos of related individuals (Parr & de Waal, 1999), or to select photos of familiar individuals based on their relationships to each other (Dasser, 1988a, 1988b). However, this is the first demonstration that animals can match photographs on the basis of a *general* social relationship concept. The photos used in the present work were of individuals that were unfamiliar to the subjects, and indeed, often depicted unfamiliar species. In some of the aforementioned studies, the subjects might have learned associations between the individuals whose photographs were presented, because they would have had experience observing these individuals interacting (Dasser, 1988a, 1988b).

In addition, the individuals depicted in the current experiment were often unrelated to each other and thus bore little "genetic" resemblance. The present findings thus go beyond demonstrating learned associations between individuals, or the recognition of perceived similarities between unfamiliar *related* individuals

(Parr et al., 2000). When subjects matched photos of related individuals, it is possible that they were using shared "genetic" features to match the photos correctly. In Experiment 2, the subjects were required to match unrelated individuals on the basis of the type of relationship that might exist between those individuals as well (e.g., mated pairs, social groups). The only information available to the subjects in these experiments was that which could be extracted from the stimuli themselves, but they could not make use of purely perceptual similarities between the stimuli. Instead they might have used several perceptual cues to form a stable abstract concept.

It is possible to argue that a series of rules were used to make the discriminations. For instance, the mother/offspring discrimination could be made by applying a rule such as "choose the photo that contains a large female and small, young individual of the same species." However this view is not necessarily incompatible with one that postulates the use of concepts. A concept may be composed of a list of features or rules that are bound together to form a single stable category or "concept," so that once these features or rules are associated together, the subject has formed a more abstract representation that encapsulates all of the individual features. The speed with which the subjects made their choices and the fact that they quickly (and sometimes immediately) reached above chance levels of accuracy supports the idea that they were able to perceive the image as representative of a concept. This is not to say that the concepts were not originally formed by attending to the presence or absence of various observable features. Whether the subjects mapped these photographs on to pre-existing concepts or learned to form new experiment-specific concepts can not be determined from these data.

Analysis of the gorilla's errors from Experiment 2 indicated that she might not have been truly extracting the concepts of social relationships from the photos but may have been simply selecting photos that included young animals. These photos may have been preferred for unknown reasons. In Experiment 1, Zuri had the most difficulty, surprisingly, with the third set of photos, which included photos of only her own species, gorillas. In examining her errors, it became obvious that she often selected photos of other juveniles even though selection of these photos was never reinforced. Because Zuri was singly housed and not able to interact with other gorillas, these juveniles may have appeared especially appealing to her. When photos of other species were presented, she was much more likely to select photos instantiating the mother/offspring relationship than to pick photos simply of young animals, suggesting that she was doing more than simply selecting photos of young animals and was actually attending to the appropriate relationship. Yet, it appears to be difficult for nonhuman primates to inhibit responding to preferred photos even when they may understand that their responses are incorrect (see also Boysen & Bernston, 1995). Zuri was given a sufficient number of sessions per day that she was quite well rewarded and may not have been highly motivated to receive a reward on every trial. It is also important to note that none of the subjects were food deprived.

It was somewhat surprising that the adult male orangutans tended to perform slightly more accurately than the gorilla on these tasks. It is sometimes expected that, because orangutans, especially adult males, lead fairly solitary lives, that they may not have developed an acute understanding of social relationships.

However, this solitary lifestyle is largely a function of the nutritional requirements of an animal as large as an adult orangutan and the patchy distribution of its food sources (Oates, 1987). The solitary lifestyle may not have driven evolution of the animals' social cognitive abilities, especially given that orangutans do congregate together when trees spawn ripe fruit. In fact, females whose territories overlap develop dominance hierarchies such that there is a definite order in which females and their young are permitted to feed from freshly fruit laden trees. In addition, juvenile orangutans often travel together until they mature (Wrangham, 1987). In a captive setting, orangutans make use of dominance information in their daily interactions and do live comfortably in groups. It seems probable that orangutans are quite sensitive to the same kinds of socially relevant information utilized by their group-living ape relatives, the chimpanzees, bonobos and gorillas.

It is also possible that the differences in performance between the subjects reflected age and not species differences. Molek was a sexually mature adult and Dinar was a subadult whereas Zuri was a young juvenile. In addition, both Molek and Dinar interact in a group of their peers. Zuri was not integrated into the gorilla troop at the Toronto Zoo at the time of these studies. Because she had been human reared she had bonded more strongly with her human caretakers than with her own species. Both of these factors may have contributed to her performance in these experiments.

It is difficult to test for abstract concepts in nonverbal species because the stimuli used are typically visual. Information about social relationships may be perceived by observing interactions between ones conspecifics. However, in these experiments, the stimuli included photos of unfamiliar individuals and unfamiliar species so that there was no prior opportunity to observe and learn associations or behavioral patterns. A simple analysis of single perceptual features of the stimuli would not have allowed the subjects to accurately perform the discriminations. It would be difficult to differentiate playing from grooming photos based on any perceptual cue and yet both subjects did this reliably. In Experiment 2, the subjects might have simply observed the number of individuals depicted in the photos and accurately matched "group" photos, and yet this discrimination was done relatively poorly. The most efficient way to perform the task would seem to be to attend to general concepts, whether these are prior existing concepts or concepts learned here through the association of complex rules and features. Therefore a tentative conclusion from these data is that both orangutans and gorillas are capable of forming abstract social concepts.

References

Boysen, S. T., & Berntson, G. G. (1989). Conspecific recognition in the chimpanzee (*Pan troglodytes*): Cardiac responses to significant others. *Journal of Comparative Psychology,* **103**, 215-220.

Boysen, S. T., & Berntson, G. G. (1995). Responses to quantity: Perceptual versus cognitive mechanisms in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behaviour Processes,* **21**, 82-86.

Call, J., Laurel, F., & de Waal, F. B. (1999). Reconciliation patterns among stumptailed macaques: a multivariate approach. *Animal Behaviour*, **58**, 165-172.

Call, J., & Tomasello, M. (1995). Use of social information in the problem solving of orangutans (*Pongo pygmaeus*) and human children (*Homo sapiens*). *Journal of Comparative Psychology,* **109**, 308-320.

Cheney, D. L., Seyfarth, R. M. (1999). Recognition of other individuals' social relationships by female baboons. *Animal Behaviour,* **58**, 67-75.

Cheney, D. L., Seyfarth, R. M., & Silk, J. B. (1995). The Responses of female baboons (*Papio cynocephalus ursinus*) to anomalous social interactions: Evidence for causal reasoning? *Journal of Comparative Psychology,* **109**, 134-141.

Cheney, D. L., Seyfarth, R. M., & Smuts, B. (1986). Social relationships and social cognition in nonhuman primates. *Science,* **234**, 1361-1366.

Coussi-Korbel, S. (1994). Learning to outwit a competitor in mangabeys (*Cercocebus torquatus torquatus*). *Journal of Comparative Psychology,* **108**, 164-171.

Dasser, V. (1988a). Mapping social concepts in monkeys. In R.W. Bryne, & A. Whiten (Eds.), *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes and humans* (pp. 85-93). New York: Oxford University Press.

Dasser, V. (1988b). A social concept in Java monkeys. *Animal Behaviour,* **36**, 225-230.

Hampton, J. A. (1981). An investigation of the nature of abstract concepts. *Memory and Cognition,* **9**, 149-156.

Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour,* **61**, 139-151.

Maestripieri, D. (1996). Primate cognition and the bared-teeth display: A Reevaulation of the concept of formal dominance. *Journal of Comparative Psychology,* **110**, 402-405.

Medin, D. L., & Smith, E. E. (1981). Strategies and classification of learning. *Journal of Experimental Psychology: Human Learning and Memory,* **7**, 241-253.

Menzel, C. R. (1997). Primates' knowledge of their natural habitat: As indicated in foraging. In R. W. Bryne, & A. Whiten (Eds.), *Machiavellian intelligence II: Extensions and Evaulations.* Cambridge, UK: Cambridge University Press.

Oates, J. F. (1987). Food distribution and foraging behaviour. In B. B. Smuts, D. L. Cheney, R. M. Seyfarth, R. W. Wrangham, & T. T. Struhsaker (Eds.), *Primate societies.* Chicago, IL: University of Chicago Press.

Parr, L. A., & DeWaal, F. B. M. (1999). Visual kin recognition in chimpanzees. *Nature,* **399**, 647-648.

Parr, L. A., Winslow, J. T., Hopkins, W. D., & DeWaal, F. B. M. (2000). Recognizing facial cues: Individual discrimination by chimpanzees (*Pan troglodytes*) and rhesus monkeys (*Macaca mulatta*). *Journal of Comparative Psychology,* **114**, 47-60.

Povinelli, D. J. (1999). Social understanding in chimpanzees: New evidence from a longitudinal approach. In P. D. Zelazo, & J. Wilde (Eds), *Developing theories of intention: Social understanding and self control* (pp. 195-225). Hillsdale, NJ: Erlbaum.

Povinelli, D. J., Nelson, K. E. & Boysen, S. T. (1992). Comprehension of role reversal in chimpanzees: evidence of empathy? *Animal Behaviour,* **43**, 633-640.

Povinelli, D. J., Parks, K. A., & Novak, M. A. (1992). Role reversal by rhesus monkeys, but no evidence of empathy. *Animal Behaviour,* **44**, 269-282.

Premack, D. (1983). The codes of man and beasts. *Behavioral and Brain Sciences,* **6**, 125-167.

Russell, C. L., Bard, K. A., & Adamson, L. B. (1997). Social referencing by young chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology,* **111**, 185-193.

Silk, J. B. (1999). Male bonnet macaques use information about third party rank relationships to recruit allies. *Animal Behaviour,* **58**, 45-51.

Spalding, T. L., & Ross, B. H. (2000). Concept learning and feature interpretation. *Memory and Cognition,* **28**, 439-451.

Tomasello, M., Hare, B., & Agnetta, B. (1999). Chimpanzees, *Pan troglodytes*, follow gaze direction geometrically. *Animal Behaviour,* **58**, 769-777.

Tversky, A. (1977). Features of similarity. *Psychological Review,* **84**, 327-351.

Visalberghi, E., Quarantotti, B. P., & Tranchida, F. (2000). Solving a cooperation task without taking into account the partner's behaviour: The case of the capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology,* **114**, 297-301.

Vonk, J. (2003a). Gorilla (*Gorilla gorilla gorilla*) and Orangutan (*Pongo abelii*) understanding of first and second order relations. *Animal Cognition,* **6** (2), March 1, 2003.

Vonk, J. (2003b). Matching based on biological categories in orangutans (*Pongo abelii*) and one gorilla (*Gorilla gorilla gorilla*). Manuscript submitted for publication.

Vonk, J., & MacDonald, S. E. (2002). Natural concept formation in a juvenile gorilla (*Gorilla gorilla gorilla*) at three levels of abstraction. *Journal of the Experimental Analysis of Behaviour*, **78**, 315-332.

Vonk, J., & MacDonald, S. E. (in press). Levels of abstraction in orangutan (*Pongo abelii*) categorization. *Journal of Comparative Psychology*.

Wrangham, R. W. (1987). Evolution of social structure. In B. B. Smuts, D. L. Cheney, R. M. Seyfarth, R. W. Wrangham, & T. T. Struhsaker (Eds.), *Primate societies.* Chicago, IL: University of Chicago Press.

Received November 25, 2002.
First revision received April 7, 2003.
Second revision received April 30, 2003.
Accepted May 2, 2003.