Performance on a Face Discrimination Task by Orangutans Reflects a Possible Interaction between Familiarity and Novelty

Catherine F. Talbot

Georgia State University

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

Recommended Citation

This Thesis is brought to you for free and open access by the Department of Psychology at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Psychology Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.
PERFORMANCE ON A FACE DISCRIMINATION TASK BY ORANGUTANS REFLECTS A POSSIBLE INTERACTION BETWEEN FAMILIARITY AND NOVELTY

by

CATHERINE F. TALBOT

Under the Direction of Sarah F. Brosnan

ABSTRACT

Faces provide humans with information on the age, sex, individual identity, and emotional state of others. Although comparatively less is known about nonhuman primates' face processing abilities, several gregarious group living species are able to discriminate conspecific faces. Here, we tested a less gregarious species, orangutans, to determine if they exhibit similar skills. Using a matching-to-sample paradigm, orangutans matched two identical portraits of unfamiliar orangutans. Next, subjects matched two different photographs of the same individual across viewpoints. During testing, subjects successfully transferred to novel photographs of familiar, but not unfamiliar, individuals with their first exposure to these stimuli. However, performance was not maintained throughout continued exposure to these stimuli, suggesting a possible novelty effect. Interestingly, orangutans performed significantly above chance when individuating familiar males, but not females. Further examination is needed to understand social organization and other social factors which were important in the evolution of face-processing.

INDEX WORDS: Individual recognition, Primates, Pongo spp
PERFORMANCE ON A FACE DISCRIMINATION TASK BY ORANGUTANS REFLECTS A POSSIBLE INTERACTION BETWEEN FAMILIARITY AND NOVELTY

by

CATHERINE F. TALBOT

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in the College of Arts and Sciences Georgia State University 2012
PERFORMANCE ON A FACE DISCRIMINATION TASK BY ORANGUTANS REFLECTS A POSSIBLE INTERACTION BETWEEN FAMILIARITY AND NOVELTY

by

CATHERINE F. TALBOT

Committee Chair: Sarah F. Brosnan

Committee: Tara Stoinski
Michael J. Owren

Electronic Version Approved:

Office of Graduate Studies
College of Arts and Sciences
Georgia State University
August 2012
ACKNOWLEDGEMENTS

I would like to thank Laura Mayo, Tara Stoinski, and Sarah Brosnan for their assistance and helpful feedback. I would also like to thank Tripp Johnston for assistance with Photoshop; Patti Ragan of the Center for Great Apes, the Little Rock and Sacramento Zoos, Bridget Wright from Zoo Atlanta and Andrea Clay for allowing me to use their photographs; Lillian Ciardelli for assistance with coding; Michael J. Beran for his advice and feedback; and the animal care staff at Zoo Atlanta for maintaining the health and wellbeing of the orangutans. CFT was funded by a National Science Foundation Graduate Research Fellowship (DGE-1051030). SFB was funded by National Science Foundation Human and Social Dynamics Grant (SES 0729244).
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .............................................................................................................. iv

LIST OF FIGURES .......................................................................................................................... vii

1 INTRODUCTION ......................................................................................................................... 1

1.1 Purpose and Expected Results of the Study ........................................................................... 4

2 METHODS ...................................................................................................................................... 5

2.1 Subjects and Housing ............................................................................................................. 5

2.2 Facial Stimuli .......................................................................................................................... 7

2.3 Apparatus and General Procedure ....................................................................................... 7

2.4 Training Procedure ............................................................................................................... 9

2.5 Shape Training ..................................................................................................................... 10

2.6 Face Training ....................................................................................................................... 10

2.7 Task Training – Individual Recognition ............................................................................. 10

2.8 Transfer Test – Individual Recognition .............................................................................. 11

2.9 Testing Procedure ............................................................................................................... 11

2.10 Data Analysis ..................................................................................................................... 13

3 RESULTS .................................................................................................................................... 13

3.1 Shape Training ..................................................................................................................... 13

3.2 Face Training ....................................................................................................................... 14

3.3 Task Training – Individual Discrimination ........................................................................ 14

3.4 Transfer Test – Individual Discrimination ........................................................................ 15
LIST OF FIGURES

Figure 1. Photograph of the experimental set-up. ................................................................. 9
Figure 2. Individual testing performance. ............................................................................. 16
Figure 3. Percentage of correct choices throughout testing. ............................................... 16
Figure 4. Performance over Time....................................................................................... 19
1 INTRODUCTION

As humans, we intuitively recognize individuals, assess their relationships with others, and perceive the disposition and intentions of others on a daily basis (Bruce and Young, 1986). In particular, faces provide us with information on the age, sex, and identity of other individuals as well as information about the emotional states of others (Tranel et al. 1988; Ekman and Oster 1979). Comparatively less is known about nonhuman primates’ abilities to discriminate and process faces. Numerous studies have found that faces are highly salient social stimuli for many animals, including nonhuman primates (Brown and Dooling 1992; Fujita 1993; Parr et al. 2000); however, there is still debate as to whether nonhuman primates process faces in the same holistic way as do humans (Bovet and Vauclair 2000; Dahl et al. 2007; Tsao and Livingstone 2008). Behavioral evidence of similar facial discrimination skills in nonhuman primates would support the hypothesis that human and nonhuman primates share similar face processing mechanisms, which presumably evolved from a common ancestor.

Many nonhuman primates are better at recognizing individuals of their own species than those of other species (Fujita 1993; Pascalis et al. 2002, Dufour et al. 2006). Moreover, both humans (Homo sapiens) and rhesus macaques (Macaca mulatta) demonstrate a preference for novel stimuli when viewing conspecific faces and objects, but do not show this preference for the faces of non-conspecific primate species (Pascalis and Bachevalier 1998). This implies that there was a strong evolutionary pressure to process conspecific faces differently, and presumably more efficiently, than the faces of other species. This would be under strong selection pressure in group living species, where individuals would do best if they recognized each other individually and remembered individuals with whom they had interacted (Pokorny and de Waal 2009). However, other evidence indicates that the ability to individuate conspecifics better than other species may merely be a result of the length of visual exposure to the different species.
For example, rhesus macaques were presented with pictures of conspecifics and domestic animals (images included full bodies). Using a habituation-dishabituation paradigm, macaques became habituated to a picture of a conspecific and then dishabituated when shown a picture of a different conspecific, suggesting discrimination between the two pictures. This effect was not found when the domestic animal stimuli were used; subjects treated domestic animals of the same species as similar, suggesting that rhesus macaques were able to individuate conspecifics, but not domestic animals. However, after several months of exposure to the domestic animals, the macaques could individually discriminate them as well (Humphrey 1974). Similarly, chimpanzees (*Pan troglodytes*) with more exposure to human faces than to other chimpanzee faces were better at discriminating human faces than they were at discriminating chimpanzee faces (Martin-Malivel and Okada 2007). In addition, chimpanzees tested on a face recognition task performed better when individuating highly familiar conspecifics across viewpoints compared to moderately familiar conspecifics (subjects had prior exposure to these individual only as test or training stimuli), and worse when individuating unfamiliar conspecifics (Parr et al. 2011). Together, these studies imply that experience and/or exposure may be a major factor in the ability to discriminate individuals.

While the studies discussed above indicate that primates exhibit a species-specific effect, they also do not necessarily provide evidence of individual recognition. One cannot rule out the possibility that subjects are able to recognize small differences in the facial stimuli of conspecifics which are not obvious (to them) in other species. Therefore, subjects may not have recognized the faces of conspecifics as belonging to particular individuals. In other words, discrimination between conspecific faces does not necessarily equate with recognition. The fact that exposure to other species increases their ability to differentiate faces supports this possibility. However, other studies provide stronger evidence that may indicate recognition. One of the most direct ways to evaluate nonhuman primates’ ability to individuate faces is to present them with a task in which they must match the same individual across different viewpoints. This task
rules out the possibility that subjects are relying on irrelevant perceptual features specific to each photograph to match the stimuli and is generally accepted as evidence for individual recognition (Parr et al. 2000; Parr et al. 2011; Pokorny and de Waal 2009; Rosenfeld and van Hoesen 1979; but see Zayan and Vauclair 1998). Using a methodology which requires a direct response from the subjects, several nonhuman primate species have demonstrated the ability to discriminate individual identities among unfamiliar conspecific faces.

An early study employed a two-choice simultaneously discrimination task in which rhesus monkeys were presented with images of unfamiliar conspecific faces (Rosenfeld and van Hoesen 1979). To control for possible stimulus cues in the individual recognition test, posture, orientation, color, size and illumination of the subject stimuli were altered. These alterations did not affect the monkeys’ ability individuate unfamiliar conspecifics (see also Bruce 1982; Heywood and Cowey 1992).

More recently, chimpanzees and rhesus monkeys were tested using a matching to sample paradigm to evaluate how quickly and accurately subjects acquired discriminations of facial images, the degree to which subjects recognized the faces of unfamiliar individuals, and which feature(s) were the most important in facial recognition. Chimpanzees and rhesus monkeys matched the faces of unfamiliar conspecifics, although rhesus monkeys required more practice to do so. Both species performed significantly worse when the eyes were masked, suggesting that the eyes are particularly important in individual recognition (Parr et al. 2000). Evidence from three species also supports this contention (Macaca mulatta: Keating and Keating, 1982; Gothard et al. 2004; Pan troglodytes: Hirata et al. 2010; and Homo sapiens: Hainline 1978), and it has recently been argued that all primates share a similar face-scanning strategy in which the eyes are of particular importance (Hirata et al. 2010). Finally, there has been a single study on New World monkeys, in capuchins (Cebus apella), showing that they, too, can discriminate both familiar and unfamiliar faces, including across different viewpoints and lighting conditions (Pokorny and de Waal 2009).
1.1 Purpose and Expected Results of the Study

Thus far, however, face discrimination skills have only been demonstrated in highly gregarious, group living species. This leaves open the question of whether all primates individuate faces in similar ways, or whether there have been specialized adaptations based on whether or not individuals spend the majority of time in social groups. In this study, we tested a far less gregarious primate species, the orangutan (*Pongo spp.*), to determine if they exhibit a similar pattern of behavior as compared to these other, more gregarious, species. Although the home ranges of orangutans overlap (te Boekhorst et al. 1990; Mitani et al. 1991) and occasionally larger aggregations form during periods of high fruit abundance (MacKinnon 1974; Rijksen 1978; Singleton et al. 2009), providing the opportunity for social encounters, orangutans spend a significantly smaller proportion of their time in groups than do other species (Galdikas 1988; Rijksen 1978). Nevertheless, orangutans may have encountered conspecifics more frequently in their evolutionary history (Meijaard et al. 2010) and currently maintain some form of social unit (van Schaik & van Hooff 1996).

Importantly, the individuals we tested came from established social groups at Zoo Atlanta, where they spend their entire day in the company of other orangutans, eliminating the possibility that any differences we might find would be due to different exposure of the individuals to other orangutans, as opposed to species differences. To test these orangutans, we used a matching to sample paradigm allowing individuals to indicate which of two possible orangutan photos depicted the same individual represented in the sample. Subjects first completed a series of training sessions before testing with photos that they had not previously seen during the individual discrimination task.

Our primary goal was to determine whether orangutans, like other more gregarious primates, can discriminate unfamiliar conspecific faces. Because orangutans share many of the same cognitive traits as chimpanzees (Russon 1998; Shumaker et al. 2001), which have demonstrated skills of individual discrimination (Parr et al. 2000; Parr et al. 2011), it is reasonable to
propose that orangutans will do similarly well. However, it has been proposed that these socio-cognitive skills evolved as a response to social complexity (Jolly 1966; Humphrey 1976; Byrne and Whiten 1988), which leads to the prediction that orangutans may not possess these skills to the extent of these other highly social species.

Second, unlike the majority of previous studies, we included stimulus sets not only of unfamiliar individuals, but of familiar individuals as well. Thus, as a secondary goal we assessed performance differences between the two categories of stimuli. On the one hand, recent evidence suggests that primates may discriminate unfamiliar individuals better than familiar individuals due to increased vigilance or a novelty effect (Mahajan et al. 2011; Pokorny and de Waal 2009), leading to the prediction that orangutans will perform better on unfamiliar trials as compared to familiar trials. On the other hand, human and chimpanzee performance decreases when discriminating unfamiliar, but not familiar individuals across viewpoints (Hill and Bruce 1996; Parr et al. 2011), and as discussed above, increased exposure may improve the ability to identify individuals (Fujita 1990; Tanaka 2003; Martin-Malivel and Okada 2007), which leads to the opposite prediction. Finally, while performance as a function of gender is rarely examined in similar studies, it may be an important social factor which influences face recognition performance. Therefore, we evaluated how secondary sexual characteristics impacted individual discriminations by examining responses between flanged males, unflanged males, and females.

2 METHODS

2.1 Subjects and Housing

The orangutans used in this study were housed in social groups at Zoo Atlanta, Atlanta, GA, USA. All subjects had indoor/outdoor access and extensive material enrichment (climbing structures, ropes and swings, barrels, and other toys). All subjects were fed their usual diet consisting of primate chow, fruits and vegetables throughout the course of the study. In addition,
feeding enrichment was provided on a daily basis as part of the husbandry routine. At no time were the subjects ever food or water deprived. Studies involved a single ape at a time. All subjects participated voluntarily, being called in from their social groups and tested in one of the indoor dens of their living area. If possible subjects were separated from other individuals to limit distractions (unweaned infants always accompanied their mothers). All procedures used in this research were approved by the Scientific Review Committee of Zoo Atlanta and the Institutional Animal Care and Use Committee of Georgia State University and were in accordance with the Association of Zoos and Aquariums animal care and management rules and the legal requirements of the United States.

Data were collected over 13 months between November, 2010 and December, 2011 on three orangutans housed at Zoo Atlanta. All 13 orangutans housed at Zoo Atlanta were presented with the experimental stimuli, but not all individuals were interested in testing or were unable to perform the basic discrimination tasks. Test subjects came from two social groups. Group 1 consisted of a 34 year old hybrid male, 28 year old Sumatran female, and two Sumatran juvenile males (ages 9 and 5), and a 1 year old Sumatran male. Test subjects included the adult female and one of the juvenile males. Group 2 consisted of all individuals from the Bornean species, including an adult male (age 18), a subadult male (age 8) and an adult female (age 19). The subadult male was the test subject.

The Sumatran female test subject was reared in a computer-enriched environment at the Georgia State University Language Research Center (Washburn et al. 2007) and participated in cognitive tasks there (e.g., Beran 2002). She and the two subadult males had previous training with a variety of cognitive tasks using the matching-to-sample paradigm on a computerized-touchscreen testing apparatus including matching social stimuli such as faces. However, the tasks presented in this study were not computerized.
2.2 Facial Stimuli

Stimuli consisted of high-quality digital color photographs of both familiar and unfamiliar individuals. Hereafter, “Unfamiliar” individuals refers to orangutans from various zoos and sanctuaries in the United States with whom subjects have never before interacted or seen, whereas “Familiar” individuals refers to other orangutans housed at Zoo Atlanta from both within their social group and from neighboring groups. All subjects had daily visual and vocal access to each of these individuals. Training stimuli used in facial tasks included 29 photographs of 14 unfamiliar conspecifics. Testing stimuli included 72 photographs of 11 familiar individuals (ranging between 5 and 40 years of age) and 62 photographs of 14 unfamiliar individuals (ranging between 6 and 41 years of age).

Using a standard graphics software package (PhotoShop CS3), photographs were cropped to only include heads and faces, making sure full flanges were visible for adult males. Stimuli included multiple photographs of males and females of all ages displaying different head positions and gaze orientations. The backgrounds of the photographs were homogenized by filling in the area around the face with solid white. Brightness and contrast were standardized to control for differences in lighting. Presentation size of the images was 12.27 cm by 17.8 cm with a resolution of 300 dots per inch.

2.3 Apparatus and General Procedure

We implemented a simultaneous matching-to-sample (MTS) paradigm (Nissen et al. 1948, Parr et al. 2000; Pokorny and de Waal 2009) in which subjects were presented with a sample image and matched one of two comparison images to the sample. The correct comparison image matched the sample on some predetermined stimulus dimension while the other image did not match. As not all subjects in this study were computer trained, we adapted this procedure to a manual task that, as in computer studies, allowed the experimenter to be completely blind to both the subjects’ options and the correct answer. To standardize the location of the
stimuli, we presented sample stimuli on a presentation board, hereafter referred to as “sample board,” while the comparison stimuli were presented simultaneously on a second presentation board, the “comparison board.” The background of the presentation board was colored to specify the MTS rule or task. For example, the presentation board for the face training had a pink background, while the board for the individual discrimination task had a black background.

To avoid experimenter cueing, we randomized, pre-sorted, and kept upside down the presentation of the stimulus sets during the session so the experimenter could access them without viewing them (and thus was blind to the correct response). For each trial, the experimenter first drew the two comparison stimuli from the pile of upside down stimuli and placed them, still upside down, on the floor, equidistant to the sample board, one to the left of the sample board and the other to the right. Neither the experimenter nor the subject could see the comparison stimuli at this point. Then the experimenter drew the upside down sample, and held it up until the subject oriented towards the image by pointing at it. The sample was fastened with Velcro onto the center of a sample board, centered in front of the subject (the experimenter still could not see the sample). Finally, the experimenter picked up the comparison board by the top, rotating it backwards so that it the orangutans could simultaneously see both comparison stimuli, while the experimenter could not. The comparison board was placed at the bottom of the sample board, with each comparison equidistant from the sample (see Figure 1). Thus, in this way the subjects saw both matches at exactly the same time, equidistant from the sample, while the experimenter could not see any of the three images. The picture the subject pointed to was accepted as their choice. Subjects were familiar with this pointing methodology, which they had used in previous studies (Stoinski, personal communication). If subjects pointed ambiguously (i.e. between the two comparison images or simultaneously pointed at both), the experimenter removed all images (still hidden from the experimenter), gave the subject a three second inter-trial interval (ITI) and restarted the trial.
The large presentation board displays the sample and the smaller comparison board displays the comparison stimuli. The comparison board allowed us to present both comparison stimuli simultaneously to the orangutans, equidistant from the sample stimulus. Note that the experimenter sat behind the display board and could not see either the sample or comparison stimuli at any time during the trial, and therefore did not know what the correct choice was. See text for more details.

Once the subject made their choice, the experimenter recorded the location of the image (left or right), and after recording this data, examined the images in order to determine whether the subject chose the matching comparison stimulus. When the correct choice was made, the experimenter verbally rewarded the ape, placed the correct choice next to and then over the sample, providing visual feedback, and then rewarded subjects appropriately. If an incorrect choice was made, the experimenter indicated this was not correct verbally as well as by placing the incorrect image up to the sample, showing that it did not match, and no reward was given to the subject.

2.4 Training Procedure

During training, a correction procedure was employed so that following an incorrect choice subjects received an ITI of three seconds and then repeated the trial. In this manner, the trial was repeated up to four times or until the subject selected the correct response, whichever occurred first (Pokorny and de Waal 2009). Only correct choices were rewarded. Training rewards consisted of cereal, flavored pellets, and Crystal Light brand juice. Subjects were given a maximum of 1 training session per day of up to 40 trials or 30 minutes. The number and dura-
tion of training sessions were recorded. Training sessions for each phase or task (shape training, face training, and task training) continued until the subjects’ performance reached a criterion of a minimum 80% success rate on two consecutive sessions (consisting of 20 trials) on two different days (as in Parr et al. 2000). Thus, the number of training sessions varied dependent upon the task or the subject being trained. This criterion was used to verify that subject understood the task before moving on to the next, more complex, task.

2.5 Shape Training

Basic simultaneous MTS training was conducted with two-dimensional colored shapes. Subjects were trained to match the identical shape to the sample shape. Shape stimuli varied on two dimensions: shape and color. Thus, subjects could use either or both of these perceptual cues to make their choice. Subjects were trained using 62 shape stimuli (31 matching pairs).

2.6 Face Training

Face training examined the ability of orangutans to match identical portraits of unfamiliar conspecifics (no photographs of familiar orangutans were used at any point during training). Subjects were trained to choose the identical comparison image that matched the sample. Face training was not designed to address whether subjects viewed these images as representations of specific individuals, but rather to provide an initial assessment of how quickly and accurately orangutans acquired the ability to discriminate complex two-dimensional facial images. Once subjects met criteria (80% success on two separate days, as described above), they proceeded to task training on the individual discrimination task.

2.7 Task Training – Individual Recognition

Using the same general method and facial stimuli, subjects were then trained on the individual discrimination task to familiarize the subjects to the task and MTS rule. Subjects were
required to match two different photographs of the same individual displaying different head positions and gaze orientations. The inclusion of this phase was important because several human studies have found a decrease in performance when the lighting, facial expression, or viewpoint of the facial stimuli were changed (Hill and Bruce 1996; Bruce et al. 1999; Bruce et al. 2001). Thus, this phase verified that subjects’ were not relying on irrelevant perceptual features of the identical photographs, such as symmetry or lighting, to match stimuli. This training phase also used only stimuli of unfamiliar individuals.

2.8 Transfer Test – Individual Recognition

Once subjects reached performance criterion (80% correct on two days; as above), they were transferred to all novel stimuli; that is, none of the photographs used in the transfer test had ever been used in any training condition and so had never been seen by the subjects. This allowed us to evaluate how well subjects recognized individuals without previous exposure to the stimuli, but after they had received sufficient experience (from Task Training) to be sure that they understood the task. To consider how individual discrimination ability might differ for familiar versus unfamiliar faces, during testing we used novel stimuli of both familiar and unfamiliar faces (note that all training involved unfamiliar individuals).

2.9 Testing Procedure

Subjects participated voluntarily, called in from their social groups by the experimenter to test separately. For testing, rewards were two sugar free colored pellets. Rewards remained consistent throughout testing. During testing, probe trials using novel photographs, which were always rewarded, were inserted among training trials (which did not involve novel photographs). Successful transfer to novel stimuli allows us to rule out performances based upon nonconceptual factors such as memorization of the training stimuli (Thompson and Oden 2000). Within each trial, the class of stimuli (familiar or unfamiliar), sex category (flanged male, unflanged male, female) and relative age were held constant so that performance could not be based on
recognizing features specific to these parameters; only the feature in question for a given task could accurately predict the correct choice. Stimulus sets for probe trials were only presented once to each subject. Subjects were never presented with images of themselves. The location (left or right) of the correct comparison stimuli was randomized; however, within a test session (and within probe trials) there were an equivalent number of correct choices located on each side.

Subjects underwent a total of 10 test sessions each with 24 trials. In each session, six probe trials of the experimental question were randomly inserted (with no more than 2 consecutive probe trials) among the 18 non-probe (training) trials. When possible, three of these probe trials used familiar stimuli and three used unfamiliar stimuli. Testing was broken down into two exposures: Initial and Continued. In the Initial Exposure, each individual was presented a single time, evaluating spontaneous categorization of individuals. Initial Exposure included four test sessions (10 familiar and 14 unfamiliar probe trials). Continued Exposure was included to provide an overall evaluation of the orangutans' facial discrimination skills. Continued Exposure consisted of 6 test sessions (18 familiar and 15 unfamiliar probe trials) that still used novel images, however individuals were presented as the sample stimulus more than once (each with a different image). Subjects received no more than one test session per day. Test sessions were given on different (and, when possible, consecutive) days.

Responses were immediately recorded on data sheets by the experimenter and test sessions were videotaped. Random number strings consisting of six digits were used to identify each stimulus presented (i.e. each individual photograph had a unique random number). Thus, the sample number never matched either of the comparison numbers. Inter-observer reliability was later conducted to verify the experimenter's accuracy in deciding whether or not the correct match was chosen. Forty-two percent of the data were recorded from the videotapes by a coder who was blind to the hypotheses to verify the experimenter data. We found perfect agreement
for all variables examined (Sample stimulus: Cohen's K = 1; left stimulus: Cohen's K = 1; right stimulus: Cohen's K = 1; Choice: Cohen's K = 1).

2.10 Data Analysis

For each test session, the experimenter recorded information on the subject, experimenter, date, session number, condition, names of the images presented as sample stimuli, the image and location (left or right) of the choice that was selected by the subject, and whether each trial was correct or incorrect. The primary dependent variable of interest was the response (correct/incorrect) and the independent variable was the condition (familiar/unfamiliar). Unfortunately, we did not have a large enough sample size to evaluate performance based on species. However, the majority of the orangutans presented as stimuli were hybrids (15 hybrids, 7 Sumatran, and 3 Bornean orangutans). We also examined sex category (flanged male, unflanged male, female) as an independent variable. We reported the number and duration of training sessions needed to reach criteria on each task for each subject. We ran a correlation between performance and test session to test if their performance changed over time, which may have indicated learning or frustration. We used binomial z scores to determine if a given number of testing trials was significantly above chance ($p < 0.05$) as well as to examine proficiency on familiar versus unfamiliar trials and between sex categories. All p-values are two tailed, except for one case in which we give both, indicated below.

3 RESULTS

3.1 Shape Training

Shape training consisted of presenting two colored, two-dimensional shapes, one of which matched the sample first presented. This was done to familiarize subjects with the matching-to-sample paradigm. Performance criterion was set at 80% correct on two consecutive ses-
essions consisting of 20 trials each. Subjects achieved criteria in an average of 17 sessions (Junior 33, Madu 5, Satu 12), or 270 trials (Junior 461, Madu 120, Satu 230).

3.2 Face Training

The face matching task presented two comparison facial stimuli, one of which was identical to the sample portrait. This task provided an initial assessment of how quickly and accurately orangutans acquired the ability to discriminate complex two-dimensional facial images. Subjects reached performance criteria in an average of 27 training sessions (Junior 36, Madu 27, Satu 20), or 350 trials (Junior 465, Madu 334, Satu 251).

3.3 Task Training – Individual Discrimination

The individual discrimination task presented two different photographs of the same unfamiliar individual displaying different head positions and gaze orientations and one photograph of a different unfamiliar conspecific, requiring the subject to match the same individual across viewpoints. This task was included to verify that subjects’ were not relying on irrelevant perceptual features of the photographs, such as symmetry or lighting, to match stimuli. Subjects met criteria in an average of 15 sessions (Junior 20, Madu 13, Satu 13), or 230 trials (Junior 312, Madu 180, Satu 199). There was no significant difference in acquisition speed between the three training conditions (shape matching, face matching, and task training; Friedman’s test: $\chi^2=4.67, P=0.097, df=2$). However, it appears that there was a slight learning effect. All of the apes reached criterion in the individual discrimination task faster than in the face matching task, but took just as long to learn the faces as they did the shapes, suggesting that they were more easily able to transfer their ability from identical to non-identical pictures of faces than from non-faces (shapes) to faces.
3.4 Transfer Test – Individual Discrimination

3.4.1 Overall Results

After reaching performance criterion of 80% correct or higher on two consecutive sessions, the orangutans were transferred to all novel photographs, which subjects had never before seen. This allowed us to evaluate how well they recognized individuals without previous experience with the stimuli, and to see whether or not they could extrapolate their knowledge of the task, learned in training, to these new photographs. Two of the three subjects (Madu and Satu) performed significantly above chance on the familiar trials, but no subject performed above chance on the unfamiliar trials (i.e. initial and continued exposure combined), despite having had the additional exposure to the same unfamiliar individuals during training, including non-probe trials during these test sessions (Binomial tests: Familiar: Junior: 61% accuracy, $Z = 0.94, p = 0.34$; Madu: 71% accuracy, $Z = 2.08, p = 0.036$; Satu: 71% accuracy, $Z = 2.08, p = 0.036$; Unfamiliar: Junior: 52% accuracy, $Z = 0.00, p = 1.00$; Madu: 62% accuracy, $Z = 1.11, p = 0.265$; Satu: 48% accuracy, $Z = 0.00, p = 1.00$, Figure 2). Analysis of the overall group performance revealed similar results (Binomial tests: Familiar: 68% accuracy, $Z = 3.16, p = 0.001$; Unfamiliar: 54% accuracy, $Z = 0.64, p = 0.520$; Figure 3).

Given these results, we also directly tested whether there was a difference in performance between the two conditions, familiar and unfamiliar. Subjects showed only a trend towards performing better with familiar faces (McNemar test: $P=0.082, N=84$). However, all subjects deviated in the same direction (see Figure 2) and, if one assumes that orangutans, like humans and chimpanzees, should perform better for familiar than unfamiliar faces, is significantly different using a one-tailed test (McNemar test: $P=0.041, N=84$; Figure 3).
Figure 2. Individual testing performance.

Individual testing performance for both familiar and unfamiliar conditions. Training trials were included to provide a baseline performance. Asterisks indicates that performance was significantly above chance level (50%), indicated by the dashed line; * p<0.05 (two-tailed); *** p<0.001 (two-tailed).

Figure 3. Percentage of correct choices throughout testing.

Percentage of correct choices throughout testing (Overall), during Initial Exposure and during Continued Exposure for both familiar and unfamiliar conditions. Training trials were included to provide a baseline performance. Asterisks indicates that performance was significantly above chance level (50%), indicated by the dashed line. Bars indicate significance between the two conditions in the overall performance (p<0.05, one-tailed) and in the Initial Exposure (p<0.05, two-tailed); *** p<0.001 (two-tailed).
### 3.4.2 Initial vs. Continued Exposure

During the *Initial Exposure* (test sessions 1-4) only one stimulus set was presented for each individual, providing a means to examine spontaneous categorization of these individuals. The results from the *Initial Exposure* trials showed a similar pattern to the overall results, with subjects successfully transferring to familiar novel stimuli with their first exposure to these individuals (80% accuracy; Binomial test: \( Z = 3.1, p = 0.001 \); Figure 3) but not unfamiliar novel stimuli (52% accuracy; Binomial test: \( Z = 0.15, p = 0.878 \); Figure 3). However, individual analyses revealed that one subject likely drove this group result. This subject (Madu) performed significantly above chance on familiar trials (Binomial tests: Familiar: 100% accuracy, \( Z = 2.85, p = 0.002 \); Unfamiliar: 50% accuracy, \( Z = 0.00, p = 1.00 \)) but the other two subjects did not, although their choices were in that direction (Binomial tests: Familiar: Junior: 70% accuracy, \( Z = 0.95, p = 0.344 \); Satu: 70% accuracy, \( Z = 0.95, p = 0.344 \); Unfamiliar: Junior: 50% accuracy, \( Z = 0.00, p = 1.00 \); Satu: 57% accuracy, \( Z = 0.27, p = 0.791 \)). As with the overall results, subjects performed significantly better on the trials depicting familiar individuals as compared to unfamiliar individuals (McNemar test: \( p = 0.022, N = 30 \); Figure 3).

*Continued Exposure* also employed all novel stimuli; however, multiple stimulus sets were presented for each individual (i.e., each individual was presented as the sample individual more than once, but individual photographic samples were never used more than once). As a group, subjects did not perform significantly above chance on either the familiar or the unfamiliar probe trials (Familiar: 61% accuracy; \( Z = 1.5, p = 0.134 \); Unfamiliar: 56% accuracy; \( Z = 0.6, p = 0.551 \); Figure 3). Individually no subject performed above chance when discriminating unfamiliar or familiar individuals (Familiar: Junior 56% accuracy, \( Z = 0.24, p = 0.815 \), Madu: 56% accuracy, \( Z = 0.24, p = 0.815 \), Satu: 72% accuracy, \( Z = 1.65, p = 0.096 \); Unfamiliar: Junior: 53% accuracy, \( Z = 0.00, p = 1.00 \), Madu: 73% accuracy, \( Z = 1.55, p = 0.118 \), Satu: 40% accuracy, \( Z = 0.52, p = 0.607 \); Figure 3). Subjects did not significantly differ in their performance on familiar and unfamiliar trials (McNemar test: \( p = 0.845, N = 45 \); Figure 3). As a side note, we emphasize that the
subjects’ performance fell on the continued trials despite the fact that performance on the non-probe test trials into which the test trials were inserted remained significantly above chance (Training: 64% accuracy; \(Z = 5.01, p < 0.0001\); Figure 3). Thus, this appears not to be a problem with attention, but some other issue related to continued exposure.

### 3.4.3 Sex Discrimination

We also examined overall performance (i.e. Initial and Continued Exposure combined) across three different sex categories (flanged males, unflanged males, and females) to explore any differences in discrimination performance. First we examined performance on all familiar probe trials. Subjects performed significantly above chance on trials portraying familiar males (Binomial tests: \(Z = 3.19, p = 0.001\); Familiar flanged males: \(Z = 2.08, p = 0.036\); Familiar unflanged males: \(Z = 2.35, p = 0.017\)), but not familiar females (Binomial test: \(Z = 0.7, p = 0.487\); there are not sufficient data points for an individual analysis). Performance did not significantly differ between flanged and unflanged males (McNemar test: \(p = 1.00, N = 23\)). Next we examined performance on unfamiliar probe trials. Not surprisingly, given that subjects did not discriminate unfamiliar individuals, group performance was not significantly above chance for any unfamiliar categories (Binomial tests: Unfamiliar flanged males: \(Z = -0.44, p = 0.664\); Unfamiliar unflanged males: \(Z = 0.91, p = 0.362\); Unfamiliar females: \(Z = 0.51, p = 0.608\)).

### 3.4.4 Change in Performance over Time

Finally, we ran a Pearson product-moment correlation to determine if learning was a possible explanation for performance. There was no correlation between performance and time over the 10 testing sessions of 24 trials each (6 probe trials each; \(R = -0.251, p = 0.485, N = 10\); Familiar: \(R = -0.577, p = 0.081, N = 10\); Unfamiliar: \(R = 0.213, p = 0.582, N = 9\); Figure 4).
Figure 4. Performance over Time.

The percentage of correct responses for familiar trials (dotted line), and unfamiliar trials (dashed line) per testing session in the individual discrimination transfer test.

4 DISCUSSION

This study is the first to demonstrate that orangutans respond to faces as salient and discriminable stimuli. During testing, two of three orangutans individuated familiar individuals, but none of the subjects performed better than chance when individuating unfamiliar individuals. Thus, like other great apes (humans: Hill and Bruce 1996; chimpanzees: Parr et al. 2011), but unlike monkeys (Pokorny and de Waal 2009), orangutans are better able to discriminate familiar than unfamiliar individuals, despite additional experimental exposure to the same unfamiliar orangutans in our design. Interestingly, as a group orangutans also performed significantly above chance when individuating familiar males, but not familiar females, and did not differ in performance when distinguishing flanged and unflanged males. This is contrary to what we anticipated, given that flanged males appear — to us — to be much more distinct, while unflanged males look very similar to females.
We note that the orangutans in this study took longer to acquire this task than has been previously reported for chimpanzees (Parr et al. 2000), but less than what has been reported for capuchin monkeys (Porkorny and de Waal 2009) and rhesus monkeys (Parr et al. 2000; Rosenfeld and van Hoesen 1979). Still, the fact that these subjects were able to accomplish this task suggests that two dimensional facial stimuli are salient and discriminable to orangutans, and they do learn to discriminate these stimuli faster than monkey species (although maybe not so rapidly as other apes; more data is needed on the two remaining ape species before we can draw this conclusion). Of course, this begs the question of whether the discrimination abilities seen in this study qualify as face recognition. Taken in the context of previous research employing similar methodologies (Parr et al. 2000; Porkorny and de Waal 2009), these results suggest that orangutans are able to recognize at least familiar conspecific faces. It is important to note, however, that while using a methodology that requires a direct response from subjects does provide strong evidence for individual discrimination, for true individual recognition to take place, subjects must not only recognize a face as familiar, but also perceive that it belongs to a specific individual. These studies provide no evidence that subjects are connecting the two dimensional facial stimuli to the actual individuals. Rather, subjects may be following a simple rule: pick the image that looks more like the sample image.

Given the extensive use of two dimensional pictures as experimental stimuli to assess a variety of cognitive and neural processes, there is surprisingly little evidence that nonhuman primates are processing these images as depictions of real-life three dimensional objects or beings. Furthermore, the studies which have attempted to provide evidence of the association of two dimensional images with real life objects have produced inconsistent results (Davenport and Rogers 1971; Winner and Ettlinger 1979; Malone et al. 1980; Martin-Malivel 1998; Judge et al. 2012). Thus although the evidence indicates that nonhuman primates respond to two dimension stimuli similarly to three dimensional objects, it is still unclear to what degree they perceive the relationship between real life objects/beings and two dimensional images. Perhaps the most
convincing evidence of true individual recognition in nonhuman primates comes from intermodal matching tasks in which subjects match individuals through two modalities (e.g. auditory and visual; Bauer and Philip 1983; Boysen 1994; Hashiya 1999; Hashiya and Kojima 2001; Adachi and Hampton 2011) or match different body parts of the same individual (Dasser 1987; Pokorny and de Waal 2008; but see also Boysen and Bernston 1989).

In accordance with Zayan and Vauclair’s (1998) hypothesis, the orangutans were better at discriminating familiar than unfamiliar faces. In our study, the subjects had already had exposure to the unfamiliar individuals, as the training stimuli consisted of the same unfamiliar conspecific individuals that were used in testing (although all novel photographs were used in test sessions). Although no training was done with familiar individuals, subjects also had a high degree of exposure to the other orangutans housed at Zoo Atlanta (familiar individuals). It is possible that the additional exposure to familiar individuals throughout the orangutans’ daily life may have aided discrimination. Previous studies have found a similar effect in humans and chimpanzees, such that performance decreases when discriminating individuals across viewpoints or lighting conditions with unfamiliar but not familiar individuals (Hill and Bruce 1996; Bruce et al. 1999; Bruce et al. 2001; Parr et al. 2011). Moreover, macaques and, more recently, chimpanzees have demonstrated the ability to match faces with other body parts of the same individual only if the individual is familiar, suggesting that real-life interactions and whole-body knowledge aided recognition (Dasser 1987; Pokorny and de Waal 2008).

While it is possible that orangutans may be better at discriminating familiar than unfamiliar individuals, we also note a second possibility that cannot be excluded. In our procedure, all training stimuli were of the unfamiliar individuals, so despite the fact that all of the testing stimuli were novel photos, subjects had been exposed to photos of these particular individuals previously in the context of this experiment. Thus, while overall they had far more exposure to the familiar individuals in the context of everyday living, they had more exposure in the experimental context to the unfamiliar individuals. Note, too, our finding that discrimination for familiar indi-
individuals was stronger on the first presentation than for any subsequent presentation, despite the use of novel photos of those individuals. Taken together, one alternate explanation for our finding that the orangutans were better at discriminating familiar than unfamiliar photos is that they are simply better at the first presentation of any individual, but become less good with successive presentations, even if successive presentations utilize novel photographs. This could be due to attentional problems (note, again, that performance on the non-probe trials did not decrease, despite the decrease in performance on probe trials). These results provide important implications for future procedural measures used to assess facial discrimination performance.

To rule out rote learning and ensure generalization, the classic methodology for categorical judgment experiments calls for the use of a transfer test employing novel stimuli which, to rule out learning based on associations between stimuli and reinforcement, are to be presented under extinctive conditions (e.g. often only the first or first and second exposure to test stimuli are used for analysis; Zayan and Vauclair 1998; Parr et al. 2000; Pokorny and de Waal 2009). We add to this that novelty effects, including both individuals and stimuli, may need to be controlled for as well. Future tests that control for these differences may help determine what, exactly, is driving orangutans’ face discrimination abilities.

We finally consider why orangutans performed better than chance on individuating familiar males, both flanged and unflanged, but not females. One possibility is that males are more salient stimuli, both to males and females. For females, males represent potential mates or threats (van Schaik 1999). Amongst males, when a flanged male meets another flanged male, the encounters are agonistic in nature, most likely because competition for mates is intense (although flanged males are also known to tolerate the presence of unflanged males; Utami et al. 2002). Moreover, males must recognize each other, as dominance hierarchies are maintained between males that regularly encounter one another (van Schaik et al. 2004). The ability to recognize and individuate familiar flanged and unflanged males may be of particular importance due to this competitive coexistence. It is also possible that orangutan male faces may simply be
more distinctive than female faces. In humans, for example, male faces are more distinctive than female faces, partly due to the pronounced nose/brow and chin/jaw area (Bruce et al. 1993). Interestingly, in one study rhesus monkeys also performed better at matching two different pictures of male monkeys compared to female monkeys (Parr et al. 2010). However, the focus of this study was to evaluate whether orangutans are able to individuate faces of other orangutans, not whether they can discriminate between the genders. More conclusive evidence on gender discrimination will need to be obtained with a study designed to directly address this question.

In the future, it will be useful to test additional species to determine the degree to which these skills vary across the non-human primates. Furthermore, orangutans’ better discrimination of familiar individuals compared to unfamiliar individuals needs further examination. It may be most advantageous for orangutans to recognize at least those orangutans in their vicinity, perhaps those that make up a loose social unit. Future comparative work examining the social factors and cognitive processes underlying facial recognition in nonhuman primates may help us better elucidate the evolution of face processing skills.

5 REFERENCES


