Effects of Two-Dimensional Noise and Feature Configuration on the Recognition of Faces In Capuchin Monkeys (*Cebus apella*)

Emily Juniper Ward
Franklin & Marshall College

Biological Foundations of Behavior: Neuroscience

BFB:490 Honors Thesis
Graduation 12 May 2007

Submitted 26 April 2007
Roger. K. R. Thompson
D. Alfred Owens
Robert Jinks
William Seeley
Michael Anderson
Eliza Reilly
Table of Contents

List of Tables.................................................................................................................................................. iii
List of Figures .................................................................................................................................................. iv
Acknowledgements......................................................................................................................................... v
Abstract........................................................................................................................................................ vi
General Introduction ...................................................................................................................................... 1
Experiment 1 .................................................................................................................................................... 6
Experiment 2 .................................................................................................................................................... 14
Experiment 3 ................................................................................................................................................... 27
General Discussion ......................................................................................................................................... 40
References....................................................................................................................................................... 47
Appendix A...................................................................................................................................................... 52
Appendix B...................................................................................................................................................... 54
Noise and Feature Configuration in Face Recognition iii

List of Tables

Table 1. Examples of Stimulus Degradation for Experiment 1.................................................................9
Table 2. Average percent of face stimuli correctly selected across all sessions at different levels of degradation.................................................................................................................................................................11
Table 3. Average grayscale values (± SD) for the surround and centers of nonhuman primate, human, and object stimuli .............................................................................................................................................................................12
Table 4. Examples of Scrambled Stimuli for Experiment 2..............................................................................18
Table 5. Global position of features for configurations of scrambled faces that subjects either rejected (i.e. correctly selecting intact face) or selected at chance level.........................................................23
Table 6. Relative positions of features for configurations of scrambled faces that subjects either rejected (i.e. correctly selecting intact face) or selected at chance level.........................................................25
Table 7. Maintenance of performance by Heiji for pairs after their initial introduction..............................33
Table 8. Correlation matrix for Heiji’s performance for human and primate pairs.................................34
Table 9. Correlation matrix for Zen’s performance for human and primate pairs. ..................................36
Table 10. Maintenance of performance by Zen for pairs after their initial introduction...........................37
List of Figures

Figure 1. Average gray-scale value distribution. The gray-scale values for individual blocks were averaged using Adobe Photoshop’s “Photomerge” to create an image representative of the average distribution of values for a) nonhumans b) humans c) objects. .................................................................12

Figures 2-4. Face Object Discrimination. Average percent of face stimuli correctly selected per session at different levels of degradation. Grayscale images of human and nonhuman primate faces and images of objects were degraded to three different resolution levels 8x8, 6x6, and 4x4. Each session consisted of 100 trials. For each trial, the subject was required to select the image of a face from a random pair of face and object images presented on a touch screen. .........................................................13

Figure 5-7. The average percent of intact pictures correctly identified across the four possible type of stimulus pairing. In the pairings, the intact (correct) picture is marked by an asterisk (*). Grayscale images of human and nonhuman primate faces were scrambled using Adobe Photoshop. For each trial, the subject was required to select the image of an intact face from a pair of images that included and intact and scrambled faces (either human or nonhuman primate). These images were randomly paired and presented on a touch screen. .................................................................21

Figure 8. Acquisition of the first pairs of stimuli by Kiki. Average percent of face stimuli correctly selected per session for the first pair of human and nonhuman primate stimuli. They grayscale image of one pair of human and nonhuman primate faces were repeatedly presented in each session. Sessions consisted of 96 trials (48 trials of the human pair; 48 trials for the nonhuman primate pair). For each trial, the subject was required to select the image of the intact face presented on a touch screen. ....... 31

Figure 9. Acquisition of the second pairs of human and primate pictures and maintenance of performance by Kiki for the first pairs. Average percent of face stimuli correctly selected per session for the first and second pair of human and nonhuman primate stimuli. They grayscale image of the pairs were repeatedly presented in each session. Sessions consisted of 96 trials (27 trials for the first primate; 24 trials for the second primate pair; 21 trials for the first human pair; 24 times for the second human). ........................................................................................................................................31

Figure 10. Acquisition of the first pairs of stimuli by Heiji. Average percent of face stimuli correctly selected per session for the first pair of human and nonhuman primate stimuli. They grayscale image of one pair of human and nonhuman primate faces were repeatedly presented in each session. Sessions consisted of 96 trials (48 trials of the human pair; 48 trials for the nonhuman primate pair). For each trial, the subject was required to select the image of the intact face presented on a touch screen. ....... 33

Figure 11. Acquisition of the first pairs of stimuli by Zen. Average percent of face stimuli correctly selected per session for the first pair of human and nonhuman primate stimuli. They grayscale image of one pair of human and nonhuman primate faces were repeatedly presented in each session. Sessions consisted of 96 trials (48 trials of the human pair; 48 trials for the nonhuman primate pair). For each trial, the subject was required to select the image of the intact face presented on a touch screen. ....... 37
Acknowledgements

This research was supported through the Charles J. G. Mayaud Award, Margery Brittain Travel Award, and the Franklin & Marshall College Marshall Scholars Award. I would like to thank the students and faculty of the psychology primate labs at Kyoto University for the generous use of their animals and facilities as well as their collaboration in this research.

I would also like to gratefully acknowledge my Oral Defense Committee as a whole for their time and effort in evaluating this thesis. but individual thanks are due as well:

Dr. Fred Owens, for reviewing this project and providing advice along the way, as well as for always bringing a smile to my face by insisting on calling me Juniper. Dr. Rob Jinks, for three years of academic advice and encouragement, and allowing me to continue to assist in teaching the Neuroscience lab even after I won the photoreceptor penetrating “contest,” 4-1. Dr. Bill Seeley, for reviewing this project and providing advice from the very beginning and for showing your support for the women’s rugby team. Dr. Mike Anderson, for being patient with my cognitive neuroscience questions and issues, for helping me search for research assistantships. Dr. Eliza Reilly, for bringing a much appreciated—and needed!—outside perspective to my committee. I have always enjoyed hearing your perspectives and of course, thanks to you, I appreciate why a urinal with “R. Mutt, 1917” signed on the side is pretty damn cool.

Several other individuals have helped keep me sane this year—or pulled up a chair next to me on my descent into madness.

Caitlin Halligan & Brittany Greulich, for bad TV shows and good documentaries (except that chicken one) that I almost without fail slept through. Along those lines, thank you, Caitlin’s couch. Sarah Bhagat, for the mutual complaining and for being a kindred soul—who-spends-too-much-darn-time-at-Whitely (R.I.P., Whitely ☮). MaryAnn Russell, without whom the psychology department—and probably most of the world—would fall apart.

Dr. Charles Heyser, for the random evenings, the D70s, for going bowling with me, and for long conversations that have made me think about many different aspects of life, the past and the future.

Finally, I would like to thank Dr. Roger K. R. Thompson, my friend and mentor, for his instrumental role in this research and in my life. Roger, I feel most deeply grateful for your innumerable contributions, your encouragement and support, and your patience with my stubbornness and anxieties. Thank you most of all for making this last year so damn memorable.

このことも言わないことも
どうもありがとうございました。
Abstract

Face recognition in Old World monkeys is remarkably similar to that of humans, but relatively little is known about this ability in New World monkeys, such as capuchins (Cebus apella). The purpose of this study was to investigate the visual information capuchin monkeys use to recognize faces. In order to study this perceptual and cognitive ability, the animals were trained to simultaneously discriminate between grayscale pictures of both humans and nonhuman primates and non-face objects. Two-dimensional visual noise was then introduced into both images to decrease face recognition (cf. Harmon & Julesz, 1973). The animals were able to make the discriminations even at a high level of image degradation, suggesting capuchins are highly sensitive to physical variations such as contrast and value distribution. In order to minimize variations in the physical properties of the stimuli, the facial features were randomly scrambled and presented with an original intact face. In this second simultaneous discrimination task, none of the animals could discriminate between the scrambled and intact face at a level significantly above chance. The capuchins acquired the ability to discriminate only after repeated presentations of the same intact and scrambled stimuli.
Effects of Two-Dimensional Noise and Feature Configuration on the Recognition of Faces in Capuchin Monkeys (*Cebus apella*)

The face is one of the most salient cues in primate social communication. Human interaction depends largely on correct perception of the faces of other to make judgments about age, gender, and emotion. Similarly, the development of larger and more complex social groups in nonhuman primates leads to less of a dependence on olfactory cues for communication and more reliance on visual communication (Andrew, 1963), such as being able to recognize individuals and species by face (Fujita, 1987; Dasser 1987).

Faces are composed of both local features (e.g. eyes, nose, and mouth) and the configural relationship between these features. Feature and configuration information, like all visual information, can be measured in terms of physical characteristics, like spatial frequency. Face information is encoded in spatial frequencies, a measure of periodic structures. High spatial frequencies (HSFs) carry information about fine edges caused by quick luminance variation; low spatial frequencies (LSFs) carry information about coarse cues caused by gradual luminance variation.

The analysis of the different spatial frequencies is an early step in visual processing (De Valois and De Valois 1988; Morrison & Schyns 2001). Visual information is processed in several parallel processing channels. One of the best-characterized of these channels is one involved in analyzing form. Form-processing differs from motion-processing even at the level of the retinal ganglion cells. P cells transmit information to parvocellular layers of the lateral geniculate nucleus of the thalamus. This information arrives at Layer 4Cβ of the striate cortex and then travels to Layer 2,3 interblobs. Many striate cells have quite narrow spatial bandwidth (Goffaux, Hault,
Michel, Vuongô, and Rossion 2005) indicating that they are sensitive to specific spatial frequencies. Information projects next to the extrastriate regions—first to the pale interstripe of the visual association cortex and then to the inferior occipitotemporal cortex of the higher-order visual association cortex. This ventral pathway is known as the “What?” pathway and is the channel for form processing.

Within the inferior occipitotemporal cortex is a specialized region dedicated to the processing of faces, known as the fusiform gyrus. In this region, there are specific cells called face cells that respond specifically to faces (Gross, Bender, & Rocha-Miranda, 1969). Furthermore, there are different patterns of brain activity when distinguishing the gender of the face, recognizing a new face, or recognizing a familiar face (Andreasen, O’Leary, Arndt, & Cizadlo, 1996) indicating that there are several different mechanisms involved in processing facial information. However, localized lesions to the fusiform gyrus results in a condition known as prosopagnosia, an inability to recognize familiar faces.

Visual information regarding faces is processed in the fusiform area. Different spatial frequencies carry different facial information. The configural cues to facial identity are mostly carried by LSFs rather than HSFs (Collishaw and Hole 2000; Leder 1996; Sergent 1986, Goffaux, et al., 2005). Spatial frequencies can be selectively displayed or blocked through filtering. Different types of filtering yields images containing spatial information that corresponds to how objects are described: existence, general form, classification, who or what it is, texture and edges (Ginsburg, 1979). Reducing the spatial frequencies in the image by as much as 28 times still leaves enough information for face identification. This suggests that image qualities can be characterized by spatial information that is relevant for a particular task, such as object recognition or face
perception. Manipulating spatial information can contribute to a better understanding of the type of information humans use to perceive objects and faces, and therefore this may be the case for capuchins.

**Two-dimensional visual noise.** Harmon and Julesz's 1973 “Lincoln Illusion” tested face recognition in humans, using pixelized or “block portraits” of Abraham Lincoln and other famous personalities. In digital images, light intensities are encoded in quantized pixel values, which are binary values assigned to analog signal amplitudes (Raff, 1993). Decreasing the resolution of an image produces sharp edges of quick contrast change. All images have random, statistical fluctuations called image or visual noise (Berbaum, Dorfman, & Madsen, 1993). Noise is extraneous information that does not transmit a signal, and can by a by-product of other processes. Noise can distort, change, or block relevant information. For example, in the context of this study, increasing the amount of quick contrast change introduces HSFs which is considered noise because they are not part of face information, but a result of degradation.

Although LSFs that are sufficient to recognize and identify faces are conserved when images are pixelized, they are masked by the increased amount of irrelevant noise in the form of HSF. This makes the faces increasingly difficult to recognize (Harmon and Julesz 1973a; 1973b). The images can be subjected to low-pass filtering which reduces the HSF visual noise to increase the recognizability of faces (low-pass filtering can be applied through computational manipulation of the images, but can also occur physiologically through squinting, which compresses the cornea. For individuals with normal vision, this slightly offsets the image on the retina, resulting in a blurred image). In low-pass filtering, visual noise is composed of frequencies higher than the frequencies used in the pixelization sampling. The LSFs remain, and the aesthetic result is a
blurring effect. Because LSFs and HSFs provide different information about faces, altering the spatial frequencies of images of faces likewise alters the ease of perceiving the face.

**Present Study.** The overall aim of this study was to investigate the extent to which low-pass filtering of degraded images would improve face recognition in capuchin monkeys. Face recognition in Old World monkeys is remarkably similar to that of humans (Lutz, Lockard, Gunderson, & Grant, 1998; Nahm, Perret, Amaral, & Albright, 1997; Wright & Roberts, 1996; Hamilton & Vermeire, 1983; Bruce, 1982; Mendelson, Haith, Goldman-Rakic, 1982; Rosenfield and Van Hoesen, 1979). However, relatively little is known about this ability in New World monkeys, which have evolved independently from Old World monkeys for approximately 35 million years (Schrago & Russo, 2003) and therefore may use different processes when responding to faces.

The goal of Experiment 1 was to test whether introducing noise into an image would decrease capuchin monkeys’ ability to simultaneously discriminate between a face and a non-face object, and if so, at what level of stimuli degradation would recognition significantly decrease to chance level. Because of the animals’ rearing environment (daily visual contact with conspecifics, both capuchin and squirrel monkeys, and humans) and because the configuration of monkey faces is similar to that of human faces (Carmel and Bentin 2002), both non-human primate faces and human faces were used as the stimuli. Previous studies have shown that Rhesus macaques employ the same oculomotor strategy when viewing monkeys and humans, suggesting that their visual system is tuned to the same features of monkey and human faces (Guo, Robertson, Mahmoodi, Tadmor, & Young, 2003).
Originally, I intended to measure the degree to which recognition at high levels of degradation could be improved by low-pass filtering the image. However, as my results show, the capuchins discriminated face from non-face even at high levels of degradation and as a result, low-pass filtering of the stimuli was not done.

I next investigated whether the animals were using variations in grayscale distribution to discriminate face from non-face objects. Hence, in Experiment 2, in order to minimize variations in the stimuli physical properties, the facial features of the face stimuli were randomly scrambled and paired with an original intact face stimulus. The goal of this second experiment was to determine if capuchins could discriminate between a configurally intact and scrambled face, as well as examine the face stimuli to see if certain scrambled configurations facilitate or impede correct discrimination.

As my results show, the subjects were unable to discriminate between intact and scrambled faces. Although, as described later, there was a slight effect of configuration on their responses, the animals did not appear to be making their responses based on holistic processing of the face. Rather, they may have been alternatively using features of the stimuli to make their choices, which is not uncommon with New World monkeys (e.g. Weiss, Kralik, and Hauser, 2001).

A final experiment was conducted to train the animals to attend to feature configuration as a basis for their choice. The goal of experiment 3 was to determine if extensive training with the same stimulus would facilitate acquisition of the discrimination task. Also, could the subjects transfer this ability to new stimuli? As the results show, even with repeated presentation of the same stimulus pair, the task of discriminating between intact and scrambled faces remains a difficult one for capuchins.
Experiment 1

The objective of experiment 1 was to determine whether the capuchins’ ability to recognize faces would be impaired in a manner consistent with the data for humans as reported by Harmon and Julesz (1973). In this study, identification and recognition of faces were impaired by the degradation through pixelization of the images. I wanted to determine to what degree capuchins monkeys can discriminate between a face and a non-face with increasing degradation of the stimuli. Because these animals had experience with operant discrimination and object choice tasks, I predicted that the capuchins would learn to discriminate between non-degraded faces and objects. I also predicted that the capuchins’ ability to discriminate between face and non-face objects would deteriorate as a function of stimuli degradation, as it had in humans. At a high level of stimuli degradation, I predicted that capuchins would not be able to discriminate between face and non-face objects at a level significantly above chance. At this point, as in the Harmon and Julesz (1973) study, I anticipated low-pass filtering should improve discrimination ability at high levels of stimuli degradation.

Method

Subjects

The subjects were three tufted capuchin monkeys (Cebus apella), named Kiki, Heiji and Zen. Kiki was pregnant at the beginning of the experiments, and gave birth at the end of Experiment 1 (the infant died shortly after birth, so Kiki was never tested with her infant). All had a variety of laboratory experiences including operant discrimination, social cognition and object choice tasks (e.g. Fujita, 2000, 2002; Fujita, Kuroshima, & Masuda, 2002; Kuroshima, Fujita, & Masuda, 2002). The animals were group housed at the Graduate School of Letters, Kyoto
University in a cage of dimensions: 300 cm (width) \( \times \) 170 cm (height) \( \times \) 62 cm (deep). They were provided with water ad libitum, and were not food-deprived, but fed at the end of the day. The housing, care, and experiments adhered to the *Guide for the Care and Use of Laboratory Primates* from the Primate Research Institute, Kyoto University.

**Apparatus**

The test apparatus was a transparent box of dimensions 45 x 45 x 45 cm. One wall of the test box had an opening 25 cm x 18 cm, behind which was a touch-sensitive 15-in CRT monitor (CDT1572-81A [Totoku, Tokyo, Japan] with Microtouch [3M, Methuem, MA]). Two levers with LEDs were located below the opening. The animals were able to retrieve food rewards (small pieces of apples or sweet potatoes) dispensed from a universal feeder (S-100; Sanso, Tokyo, Japan) to a cup on the left side of the apparatus. A hand-assembled computer (CPU: K6II 400 MHz [Advanced Micro Devices, Sunnyvale, CA]) in a separate room controlled the apparatus using Microsoft Visual Basic 5.0. Kiki and Heiji were tested individually in a lit testing room, while Zen was tested in a darkened testing room. This was based solely on their original individual training conditions with the apparatus. White nose masked external noises in both rooms.

**Stimuli**

Fifty grayscale pictures of human faces (26 female) (24 male) and 50 grayscale pictures of nonhuman primate faces (various species) were acquired from public stock image websites including Stock.XCHNG (Table A1-A2). The faces were oriented forward. The images were cropped to include key facial features—many included most of the head as well, but as little background as possible to eliminate salient background cues in the pictures.
One hundred grayscale pictures of various objects (Table A3) were also selected. The pictures of objects were cropped as close to object as possible. Because many of the objects were photographed on a white background, I determined the average gray value of the image using the “Average” filter in Adobe Photoshop CS and replaced the white with this value. This reduced the contrast between image and background. To equate luminosity differences, the images of the faces and objects were averaged and the gray scale value was recorded. The average value for the objects was significantly higher than that of the faces. The images of the objects were made 65% darker with Adobe Photoshop CS, such that there was no significant difference in brightness between the objects and the faces. All images were scaled to 160 x 160 pixels.

Procedure

Training. The monkeys had been trained to use the touch screen program by first holding down the lever on either the bottom right or left side of the monitor before making a selection with the other hand. Kiki used the left lever, while Heiji and Zen used the right lever. After the subject had held the level for 1 s, an image of an object and a face were presented to the monkey on the screen. The objects were randomly paired with a face (either human or nonhuman primate) and appeared randomly in 10 different locations on the screen. The subject made a selection by pressing one image on the screen with one hand while holding down the level with the other (to prevent the animal from using both hands). The animals were trained to simultaneously discriminate between the two images. If the face was selected, a light turned on and bell sounded, and the animal was rewarded with a food reward. An incorrect choice yielded a different sound, and no light or reward. There was then a 3 s delay before the animal could press the lever to introduce the next selection. In each session, the animal made 100 selections. Training criterion was 90% correct for two
consecutive sessions. The sessions were conducted between 11:00 and 14:00 in June and July 2006 at Kyoto University, Kyoto, Japan.

**Testing.** Once all three reached criterion and moved onto the testing phase, the animals participated in five testing sessions, with 100 trials per session. Forty trials were with non-degraded stimuli, the same as training. This was to ensure that the animal could make correct discriminations and be rewarded. Twenty of the remaining trials were with 8x8 blocks, twenty with 6x6 blocks, and twenty with 4x4 blocks, intermixed randomly within the session (Table I). All images were degraded with the ‘mosaic’ filter in Adobe Photoshop CS, changing the ‘cell size’ to achieve desired degradation level.

<table>
<thead>
<tr>
<th>Examples of Stimulus Degradation for Experiment 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Human</strong></td>
</tr>
<tr>
<td>Non-degraded</td>
</tr>
<tr>
<td>8x8</td>
</tr>
<tr>
<td>6x6</td>
</tr>
<tr>
<td>4x4</td>
</tr>
</tbody>
</table>
Objects were paired randomly with either a human face or nonhuman primate face at the same level of degradation and were presented randomly in ten different locations on the screen. If the face was selected, a light turned on and bell sounded, and the animal was rewarded. An incorrect choice yielded a different sound, and no light or reward. There was then a 3 s delay before the animal could press the lever to introduce the next selection. In each session, the animal made 100 selections. The sessions were conducted between 11:00 and 14:00 in June and July 2006 at Kyoto University, Kyoto, Japan.

Results

*Face-Object Discrimination.* The average percent correct per session for each animal is displayed in Figures 1-3. All three animals discriminated face stimuli from non-face stimuli significantly above chance at all levels of degradation (Table 2). Since the animals were presented with two pictures between which to make a selection, chance level was 50% correct.

Overall, Kiki correctly discriminated between non-degraded pictures significantly better than those at the 4x4 level (repeated measures ANOVA, $F_{3,16}=5.710; P=0.007$). No significant effect of degradation except from intact to 4x4 level (Figure 2).

Heiji correctly discriminated between the non-degraded pictures significantly better than images degraded to 8x8 ($F_{3,16}= 4.239, P=0.022$), 6x6 ($F_{3,16}= 8.279, P=0.001$) and 4x4 ($F_{3,16}=9.830, P<0.001$). He correctly discriminated between images at 8x8 significantly better than at 4x4 ($F_{3,16}= 3.510, P=0.040$) (Figure 3).

Zen also correctly discriminated between the non-degraded pictures significantly better than when the images were degraded (for 8x8, $P<0.001$), 6x6, ($P=0.006$) and 4x4, ($P=0.002$).
However there was no effect of increasing degradation. That is, 8x8, 6x6, and 4x4 were statistically the same (Figure 4).

Table 2.

Average percent of face stimuli correctly selected across all sessions at different levels of degradation.

<table>
<thead>
<tr>
<th>Degradation Level</th>
<th>Kiki</th>
<th>Heiji</th>
<th>Zen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-degraded</td>
<td>100</td>
<td>99</td>
<td>90</td>
</tr>
<tr>
<td>8x8</td>
<td>92</td>
<td>82</td>
<td>63</td>
</tr>
<tr>
<td>6x6</td>
<td>87</td>
<td>72</td>
<td>71</td>
</tr>
<tr>
<td>4x4</td>
<td>78</td>
<td>67</td>
<td>68</td>
</tr>
</tbody>
</table>

Note: Binominal P≤0.01, except for a, where P=0.012

Stimuli Analysis. The stimuli of highest degradation (4x4) were analyzed. Figure 1 shows a representation of the average value distribution for the three stimulus types. The RGB values for the 12 surround blocks were averaged and compared to RGB values for the four averaged center blocks for all images using an independent t-test. Colors can be represented by RGB values, which give the amount of red, green, and blue in the color. For grayscale images, the amount of red, green, and blue is equal, and can therefore a grayscale value be represented by only one value. Note that black is 0 and white is 255. A summary of the values for the center and surround for the three stimulus types is provided in Table 3.
Figure 1. Average grayscale value distribution. The grayscale values for individual blocks were averaged using Adobe Photoshop’s “Photomerge” to create an image representative of the average distribution of values for a) nonhuman primates b) humans c) objects.

Table 3. Average grayscale values (± SD) for the surround and centers of nonhuman primate, human, and object stimuli

<table>
<thead>
<tr>
<th>Nonhuman Primate</th>
<th>Human</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surround</td>
<td>Center</td>
<td>Surround</td>
</tr>
<tr>
<td>121±28</td>
<td>115±24</td>
<td>110±48</td>
</tr>
</tbody>
</table>

The degraded nonhuman primate faces are homogenous, with no significant center and surround contrast change ($t_{98,0.05}=1.068; \ P=0.288$). On average, the degraded human faces have a significant dark-surround/light-center contrast ($t_{98,0.05}=-4.171; \ P<0.001$). Conversely, the degraded object stimuli tend to have a significant light-surround/dark-center contrast ($t_{98,0.05}=10.165; \ P<0.001$). The average difference in contrast between the center and surround is greatest in the object stimuli, 48 shades different. In human faces, there is almost an equal difference, -35 shades (the negative sign indicates that the center was lighter than the surround). The difference between center and surround in the nonhuman primate faces is only 6 shades. The average contrast difference between the center and the surround of the stimuli differed significantly between all three types ($F=95.658; \ P<0.001$).
Figures 2-4. *Face Object Discrimination.* Average percent of face stimuli correctly selected per session at different levels of degradation. Grayscale images of human and nonhuman primate faces and images of objects were degraded to three different resolution levels 8x8, 6x6, and 4x4. Each session consisted of 100 trials. For each trial, the subject was required to select the image of a face from a random pair of face and object images presented on a touch screen.
Discussion

The results in Experiment 1 reveal that using the current set of stimuli at all levels of degradation chosen for this experiment, capuchins can correctly discriminate between a face and an object at a level significantly above chance. While I predicted they would be able to do this at low levels of degradation, their above-chance performance at high levels of degradation was inconsistent with my hypothesis. Because of the low level of acute visual information in the degraded stimuli, my results raise the question of whether the animals perceive the images as faces or are attending to some other physical characteristic of the stimuli, such as contrast and/or distribution of values, in order to make the correct discrimination. To investigate this, the average grayscale distributions of all the 4x4 images from the three categories of stimuli (Figure 1) were produced using Photoshop’s photomerge function, revealing significant differences between the center and the surround.

While all three monkeys maintained a high percent correct for non-degraded images, their performance for the three levels of degradation show differences that may indicate the formation of discrete categories, based the significant differences in distribution and contrast between the three types of stimuli. Since these differences exist, the capuchins may have been using distribution and contrast differences to make their discriminations.

Experiment 2

Since the degraded stimuli are composed of significantly different value distributions, I wanted to minimize variations in the physical properties of the stimuli to control for the varying value distributions present in Experiment 1. One way to accomplish this was to keep the same components in both the face and non-face images. To achieve this, the facial features of the face
stimuli were randomly scrambled creating a non-face chimera that retained all the same physical features (and thus grayscale values) as the original intact face (see Appendix B). Although these images are perceived as grotesque distortions by humans, studies suggest that inverted and scrambled face images are actually less salient to nonhuman primates than upright, intact faces, as evidenced by reduced viewing times and reduced number of fixation for the manipulated images (Guo, Robertson, Mahmoodi, Tadmor, and Young, 2003; Neiworth, Hassett, Sylvester, 2007).

In humans, the configural relationship between facial features is a critical component in facial recognition. For familiar faces, humans use information about the configuration of internal features (such as eyes, nose, mouth, etc) for recognition (Ellis, Shepherd, & Davies, 1979; Young, Hay, McWeeny, Flude, & Ellis, 1985; Rhodes, 1993). Babies, for example, spend about twice as long looking at a normal face than a scrambled face (Frantz, 1961). This suggests that there is an innate ability to recognize faces that is dependant on normal feature configuration. Furthermore, there is evidence for perceptual specialization for normal feature configuration. Webster and MacLin (1999) investigated a profound featural aftereffect for adaptation to distorted faces, where participants reported that normal facial feature configuration appeared distorted (e.g. expanded) after viewing an image of a face that had been distorted to the other extreme (e.g. compressed). Therefore, there appears to be a generic face template against which distortion and manipulation can be judged.

However, within normal feature configuration, research has shown that the eyes and the mouth are especially used in processing facial information (Thompson, 1980). Rhesus monkeys exhibit similar eye scan patterns, focusing on the eyes especially, for familiar and unfamiliar monkey faces, as well as monkey and human faces. However, the probability of refixation on the
eye regions is greater for unfamiliar faces, which suggests that the eyes are important for initial encoding of identity (Guo, et al, 2003).

Another example of the importance of specific features is illustrated by an illusion known as the Margaret Thatcher Illusion. Here, when the eyes and mouth of an image of the former prime minister are turned upside down, the face appears strikingly grotesque. Yet, when the whole image is turned up-side down, the distortion is less salient. Furthermore, when an entire image of an intact face is inverted, our ability to recognize the face becomes impaired. This impairment is known as the face-inversion effect, and is evidence for specialized neurological face processing (Diamond and Carey, 1986; Farah, Tanaka, & Drain, 1995).

Past studies have shown that the brain responds differently to intact faces and faces in which the features have been manipulated. For example, ERP studies with humans have shown the latency of the N170 component which is thought to represent the activation of brain processes related to face perception is delayed when viewing inverted faces of isolated facial features (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Taylor, Itier, Allison, Edmonds, 2001; Rossion and Gauthier, 2002). A similar effect has also been found in macaques (Pineda and Nava, 1993).

Behaviorally, in macaques, when a face is inverted or scrambled, viewing duration and number of fixations per image decreases. However, even when displaced, the eye region in these faces remains the primary area of fixation (Guo, et al., 2003). Although the frequency of refixation decreases from intact face to scrambled or inverted, the eyes remain the most salient feature regardless of the arrangement of the other features. The extent of the saliency may depend on the global arrangement of facial features. For example, the eyes in the scrambled images were much less salient than in either the upright or the inverted images.
Since the animals easily discrimination faces from objects, I predicted that they could transfer this ability to the task to discriminate a configurally-intact face from a face in which the features were randomly scrambled. Having reduced physical variations between the two presented images by scrambling the features of the face, I hypothesized that the capuchins' ability to discriminate between intact face and scrambled face would deteriorate as a function of stimuli degradation, consistent with the findings of Harmon and Julesz (1973a; 1973b). Furthermore, the capuchins would not be able to discriminate between an intact face and scrambled face at a level significantly above chance at a high level of stimuli degradation. Since the physical variations were minimized by using the scrambled face stimuli, the animals would be less likely to use physical differences as a criterion for making their selection. I predicted that low-pass filtering would improve discrimination ability at high levels of stimuli degradation.

**Method**

*Subjects and Apparatus.*

The same subjects and apparatuses from Experiment 1 were used in Experiment 2.

*Stimuli*

The human and nonhuman primate face images from Experiment 1 were scrambled by hand using Adobe Photoshop (Table 4 and Appendix Table B1). To scramble the human faces, the facial features were classified by location, where the right eyebrow was assigned the number 1; right eye assigned 2; left eyebrow, 3; left eye, 4; nose, 5 and; mouth, 6. Right and left were designated from the perspective of the subject in the image. Therefore, an intact face was numerically coded, “123456,” where the numbers represent feature, and the position of the numbers in the string represent the location of the features on the face. Using these designations,
these features were randomly assigned a new location using an Excel spreadsheet. For example, 6 could be shifted to the 4 position, meaning that the mouth would be moved to the position of the right eye, such as in the string “132645,” corresponding to the human in Table 4. There was a minimum three-unique-feature scramble necessary. That is, simply switching the eyes was not sufficient to be considered scrambled.

Table 4.  

<table>
<thead>
<tr>
<th></th>
<th>Human</th>
<th>Nonhuman Primate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intact</strong></td>
<td><img src="image1" alt="Intact Human" /></td>
<td><img src="image2" alt="Intact Nonhuman Primate" /></td>
</tr>
<tr>
<td><strong>Scrambled</strong></td>
<td><img src="image3" alt="Scrambled Human" /></td>
<td><img src="image4" alt="Scrambled Nonhuman Primate" /></td>
</tr>
</tbody>
</table>

Nonhuman primate features were also classified by location, but since nonhuman primate generally lack distinctive eyebrows, there were two less features used to scramble the image. The right eye corresponded to number 1; left eye to 2; the nose to 3; and the mouth to 4. Because there were two less features, there was a minimum two-unique-feature scramble necessary (Table 4 and Table Appendix B2).

In Adobe Photoshop, the features were moved and blended by hand to remove harsh edges that introduce visual noise into the image.
Procedure

Training. Each session again consisted of 100 trials. Intact faces were paired randomly resulting in four possible pairings 1) Intact Human-Scrambled Human, 2) Intact Human-Scrambled Nonhuman Primate, 3) Intact Nonhuman Primate – Scrambled Human and 4) Intact Nonhuman Primate – Scrambled Nonhuman Primate. The experimentally correct response was selecting the intact face, whether human or nonhuman primate. These images were not degraded. They were presented randomly on the screen, and the training procedure was otherwise the same as that in Experiment 1. Criterion was two successive performances at 90% to advance to the testing (degradation) phase.

Results

None of the animals discriminated between scrambled and intact faces at a level significantly above chance (50%), and therefore did not progress to the degradation testing phase. A one-way ANOVA revealed there was also no significant difference between the three animals ($F_{2,80}=0.244, P=0.784$) in their average percent correct for the four different stimuli pairings. Individually however, they showed significant differences between the four possible types of stimuli pairing.

A repeated measures ANOVA revealed a significant difference between the percent Kiki correctly identified in the Intact Human – Scrambled Nonhuman Primate pairing and the Intact Nonhuman Primate – Scrambled Human pairing ($F_{3,32}=1.874, P=0.045$). In other words, Kiki incorrectly chose scrambled human when paired with an intact nonhuman primate, but correctly chose intact human when paired with scrambled nonhuman primate (Figure 5).
Heiji also incorrectly chose scrambled human when paired with an intact nonhuman primate but correctly chose intact human when paired with scrambled nonhuman primate ($F_{3, 32} = 8.850, P < 0.001$). Unlike Kiki, he incorrectly chose scrambled human paired with intact nonhuman primate significantly more frequently than when it was paired with an intact human ($F_{3, 32} = 7.749, P < 0.001$). He correctly identified intact humans significantly more often than intact nonhuman primates when paired with scrambled nonhuman primates ($F_{3, 32} = 3.466, P = 0.027$). He also incorrectly chose the scrambled human when paired with an intact nonhuman primate significantly more often than incorrectly choosing the scrambled nonhuman primate ($F_{3, 32} = 4.512, P = 0.009$) (Figure 6).

Zen also incorrectly chose scrambled human when paired with an intact nonhuman primate but correctly chose intact human when paired with scrambled nonhuman primate ($F_{3, 32} = 10.197, P < 0.001$). He incorrectly chose scrambled human paired with intact nonhuman primate significantly more frequently than when it was paired with an intact human ($F_{3, 32} = 8.751, P < 0.001$). He correctly identified intact humans significantly more often than intact nonhuman primates when paired with scrambled nonhuman primates ($F_{3, 32} = 5.912, P = 0.002$). He also incorrectly chose the scrambled human when paired with an intact nonhuman primate significantly more often than incorrectly choosing the scrambled nonhuman primate ($F_{3, 32} = 4.624, P = 0.008$) (Figure 7). Unlike Heiji and Kiki, he correctly identified an intact human from a scrambled human significantly more often than a nonhuman primate from a scrambled nonhuman primate.
Figure 5. Kiki

Figure 6. Heiji

Figure 7. Zen

Figure 5-7. The average percent of intact pictures correctly identified across the four possible type of stimulus pairing. In the pairings, the intact (correct) picture is marked by an asterisk (*). Grayscale images of human and nonhuman primate faces were scrambled using Adobe Photoshop. For each trial, the subject was required to select the image of an intact face from a pair of images that included and intact and scrambled faces (either human or nonhuman primate). These images were randomly paired and presented on a touch screen.
Error Analysis

In order to investigate the effect of feature configuration on the subjects’ ability to discriminate between scrambled and intact faces, all scrambled faces were assigned the numerical string used to scramble the features where the numbers represent features, and the position of the numbers in the string represent the location of the features on the face (to reiterate, for human photographs, “1” corresponded to the left eye-brow, “2” corresponded to the right eye-brow, “3” corresponded to the left eye, “4” corresponded to the right eye, “5” corresponded to the nose, and “6 corresponded to the mouth.)

Due to the bilateral symmetry of faces, both eye-brows and eyes were considered equivalent features. Thus, the total number of distinct features was simplified for analysis. A normal face was now coded as “113356” and a scrambled face would be coded, for example, “631513.” Nonhuman primate stimuli were dealt with in the same way. Therefore, intact non-human primate faces were coded, “1134” (“1234” before simplification) and scrambled faces coded, for example, “4113” (“4213” before simplification).

After the scrambled faces were coded in this manner, there were 41 unique configurations in the 50-image human stimuli; 10 unique configurations in the 50-image non-human primate stimuli.

Binomial tests were run on the frequencies of selection for each unique configuration to determine if the number of times the animals correctly chose the intact face significantly differed from chance. For example, for the human stimuli, two faces had configuration “163513”. Both faces appeared once in each session (totaling two presentations of configuration “163513”). Therefore, after nine sessions, configuration “163513” had been presented to the animal 18 times.
If he correctly rejected the image (thereby selecting the experimentally correct intact face) more than nine or 10 times, then his performance was above chance. If he selected the scrambled configuration at or around nine times, then this configuration had no effect and was selected at chance.

Of the 41 unique configurations in the human stimuli, for 17 configurations, the animals rejected the scrambled at a level significantly above chance. For the remaining 24 configurations, the animals selected at chance between the scrambled and intact face. The configurations were summarized using two criteria: the global position and relative position of the eyes and the mouth and nose. The eyebrows were not included in the data analysis because in many of the images they were not prominent features and served as placeholders when the faces were scrambled.

The global position criterion was used to classify the configurations based on whether the features were separated on different halves (top and bottom) or different sides (left and right). If one of the features in question appeared on the midline (at position 5 [nose] or 6 [mouth]), it was included in the “different sides” group, but was not included in that group when both features appeared on the midline. The global position of features for the 17 configurations selected against and 24 configurations selected at chance is summarized in Table 5.

Table 5.

Global position of features for configurations of scrambled faces that subjects either rejected (i.e. correctly selecting intact face) or selected at chance level.

<table>
<thead>
<tr>
<th>Percent of Faces with Configuration (%)</th>
<th>Eyes</th>
<th>Mouth &amp; Nose</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correctly Rejected</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Different Halves</td>
<td>82</td>
<td>29*</td>
</tr>
<tr>
<td>Different Sides</td>
<td>94</td>
<td>65*</td>
</tr>
<tr>
<td>Selected at Chance</td>
<td>67</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>88</td>
<td>92*</td>
</tr>
</tbody>
</table>

Note: Asterisks denotes significance (binomial P<0.05)
Binomial tests were used to determine if the number of rejected faces with a particular configuration differed from the number of faces with the same configuration that were chosen at chance level. The hypothesized values for the test were the percentage of faces selected at chance level that had the given configurations (Table 5).

There was no effect of the position of the eyes, whether on different halves (binomial, $P=0.129$) or on different sides (binomial, $P=0.354$). However, there was a significant effect for the position of the mouth and nose. Scrambled faces that were correctly rejected had the mouth and the nose on the same half more frequently than scrambled faces that were chose at chance level (binomial, $P=0.015$). Furthermore, scrambled faces that were correctly rejected also had the mouth and the nose on the same side (binomial, $P=0.002$).

The position of the features relative to each other was also investigated. For this analysis, features could be separated by zero- through three-degrees of separation. To illustrate this, when two features are located $0^\circ$ from each other, they are horizontally adjacent (e.g. for an intact face, eye separation is $0^\circ$). When the features are separated by $1^\circ$, they are vertically adjacent (e.g. for an intact face, there is $1^\circ$ of separation between the mouth and nose). For $2^\circ$, the features are separated by a middle feature (e.g. the mouth and eyes are separated by $2^\circ$, with the nose between). Finally, features separated by $3^\circ$ have two features between them (e.g. the mouth and eyebrows are separated by $3^\circ$, with the nose and one eye between).

The 17 rejected configurations and the 24 configurations selected at chance were analyzed and put into groups based on feature separation. The total number of faces characterized by each particular separation level was recorded and converted to a percentage. The percentage of faces
rejected with each feature separation level was compared to the percentage of faces selected at chance with the corresponding feature separation. Binomial tests were used to determine if the two differed in their characteristic feature separation (Table 6).

Table 6.

Relative positions of features for configurations of scrambled faces that subjects either rejected (i.e. correctly selecting intact face) or selected at chance level.

<table>
<thead>
<tr>
<th>Percent of Scrambled Faces with Configuration (%)</th>
<th>Eye Separation</th>
<th>Mouth &amp; Nose Separation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correctly Rejected</td>
<td>0°</td>
<td>1°</td>
</tr>
<tr>
<td>12</td>
<td>18*</td>
<td>47</td>
</tr>
<tr>
<td>Selected at Chance</td>
<td>8</td>
<td>42</td>
</tr>
</tbody>
</table>

Note: Asterisks denotes significance (binomial P<0.05)

There were fewer rejected faces that had configurations with 1° of eye separation, (binomial, P= 0.035), but there was no difference for the other degrees of eye separation (0: binomial, P=0.418; 2: P=0.091; 3: P=0.313). Alternatively, there were more rejected faces that had configurations with the mouth and the nose separated by 1° (binomial, P=0.010), but fewer with 2° separation (binomial, P= 0.050), compared to chance level faces. There was no difference for configurations with 0° and 3° separation (P=0.582; P=0.282, respectively) (Table 6).

Discussion

Contrary to my original hypothesis, the capuchins failed to discriminate intact faces from scrambled faces at a level significantly above chance. Furthermore, my results suggest that the animals chose human faces more frequently than nonhuman primate faces, even if the human face was scrambled, and thus the procedurally incorrect choice. That is, scrambled and intact human faces were perceptually equivalent to a greater extent than scrambled and intact nonhuman
primate faces. This may be an effect of the capuchins’ environment: a laboratory setting with continuous human interaction. Past research has shown that the face cells of laboratory-reared macaques respond more to human faces than monkey faces (Mikami, Nakamura, & Kubota, 1994). This suggests that experience in discriminating visually similar objects of a novel category could lead to more neurons being devoted to this category.

Furthermore, the eyes did not seem to facilitate correct identification. There was no significant effect of the global position of the eyes. The only difference between the configurations that were rejected and those that were selected at random was that the eyes were separated by 1° in fewer faces for the correctly rejected configurations.

However, there were significant effects of both the global and relative positions of the mouth and the nose. The faces that were rejected did not have the mouth and nose on different sides or different halves as often as faces selected at chance, suggesting that close global proximity of these two features assists with discrimination. Furthermore, more scrambled faces were rejected if the nose and mouth were separated by 1° than those selected at chance. However, there was no significant difference between the two groups for 0° separation of the mouth and nose. When the degree of separation increased to 2°, fewer faces were rejected with this configuration compared to chance. This result provides more support that close proximity of the nose and mouth facilitated correct discrimination.

Although there was no direct effect of the position of the eyes on selection, perhaps having the nose and mouth in close proximity permits the animals to attend to positional shifts of the eyes with more accuracy. If this is the case, there may be no prevalent pattern of eye position that
helps the animals discriminate (which is consistent with my findings), but eyes may still be a salient characteristic for the discrimination task.

Overall, these results suggest that the behavior of all three monkeys is controlled by something other than overall configuration to make their discriminations. It is common for monkeys to attend to alternative features to solve tasks such as these. For example, prior studies that have failed to find an inversion effect in primates may have been due to a lack of constraints of experimental methods. Weiss, Kralik, and Hauser, (2001) studied face processing in cotton-top tamarins, and found no inversion effect. However, more detailed analysis revealed that the tamarins were in fact attending to distinctive features and their location as a basis for their response. The animals’ reliance on alternative methods for making discriminations can lead to a convoluted understanding of the potential mechanisms of face processing in nonhuman primates.

Experiment 3

Because all animals performed at chance level in Experiment 2, a third experiment was conducted to determine if the animals could be trained to attend to feature configuration to make their discriminations. This experiment was conducted at Kyoto University as an extension of Experiment 2, but not with my direct consultation. In this final experiment, the stimuli presentation was simplified so that the animals were repeatedly presented with the same stimulus pair. Since the animals would see the same stimulus and would immediately know that their previous response (and therefore criteria for discrimination) was experimentally incorrect, they would have the opportunity to develop new selection techniques. Through these means, an attempt was made to train the animals to respond to configuration of the features, rather than experimentally extraneous information.
After the subjects reached criterion with the first pairs of repeatedly presented stimuli, the next pairs would be introduced as a transfer test. When nonhuman primates are trained on a long series of discrimination problems, they can quickly solve new problems that are introduced (Harlow, 1949). Thus, I predicted that if the animals learned to attend to scrambled configuration, they should therefore be able to transfer this ability to the new pair of images. Because of the complexity of the images and the difficulty that the animals had had in Experiment 2, it was predicted that initial performance in this test would be at chance level, but over the sessions all three animals should have been able to acquire the necessary ability to make a proper discrimination (i.e., the ability to discriminate intact from scrambled based on configuration should transfer). Furthermore, the acquisition of the subsequent pairs of stimuli would occur more quickly than the initial pairs.

Method

Subjects, Apparatus, Stimuli

The same subjects, apparatus, and stimuli from Experiments 1 and 2 were used in Experiment 3.

Procedure

96-trial sessions were ran in which the first scrambled/intact pair of human and nonhuman primate pictures were presented repeatedly. Within each session, the human pair and the nonhuman primate pair was presented 48 times each, alternating randomly. The animal was again required to choose the intact face over the scrambled face. Criterion was always set at 90% correct overall.
When the subject reached criterion of 90% correct on one session, a second pair of human pictures and nonhuman primate pictures was introduced. Within the 96 trials, the first nonhuman primate pair was presented 27 times; second nonhuman primate pair was presented 24 times; the first human pair was presented 21 times; the second human pair was presented 24 times.

When the subject reached criterion, the third pairs were introduced. Within the 96 trials, the first pairs were presented 12 times each; second pairs was presented 12 times each; and the third pairs were presented 24 times each. When the fourth pairs were introduced, the first pairs were presented eight times each; second pairs was presented eight times each; the third pairs were presented eight times each, and the fourth pairs were presented 24 times each. For the fifth pair, the first, second, third, and fourth pairs were presented six times each, and the fifth pairs were presented 24 times each. For the sixth pairs, the first pairs were presented eight times each, the second through fifth pairs were presented seven times each, and the sixth pair was presented 12 times.

Results

For the first pairs, Kiki reached overall criterion after 37 sessions (3552 trials) (Figure 8). She correctly chose the intact human face significantly more than she chose the intact primate face (paired $t_{36,0.05} = -10.496; P<0.001$). In fact, she never reached criterion for discriminating a scrambled monkey face from an intact monkey face (up to 83 % correct in session 35). A correlation of the data revealed that Kiki’s performance for the first human pair and first primate pair were significantly related, $r= 0.560, n=51 P<0.001$, (two tail).

After reaching criterion, the second scrambled/intact pairs of human and non-human primate pictures were introduced. Overall criterion was 90% correct performance for all pairs (first
human and primate pair and second human and primate pair). Kiki reached overall criterion after 14 sessions. She reached criterion for acquisition of the second pair of monkey faces (92% correct in session 7), but peaked at 88% correct for the second pair of human faces (session 12). Her performance for the primate pair was significantly better than her performance for the human pair (paired $t_{13,0.05}=4.502$, $P=0.001$). However, her maintenance performance for the first pair of primate faces fluctuated below criterion (80±9% correct), while maintenance performance for the first pair of human faces remained significantly higher (paired $t_{13,0.05}=-5.576$; $P<0.001$) at 97±5% correct (Figure 9). Because the three animals were each tested daily, and Kiki required a greater number of sessions to reach criterion for the first pair, she has not yet been tested with the third pairs of faces.

Heiji required only 12 sessions to reach overall criterion for the first pairs (1152 trials) (Figure 10). His performance was significantly better for the human pair than the primate pair (paired $t_{11,0.05}=-3.513$; $P=0.005$); he reached criterion for the human pair after eight sessions and reached criterion for the primate pair after nine sessions. There was a correlation between Heiji’s performance for the first human pair and his performance for the first primate pair, $r=0.634$, $n=45$, $P<0.001$. His performance for the first human pair was also correlated with his performance for several subsequent pairs, as shown in Table 8. There were fewer correlations with his performance for the first primate pair.

For the second pairs, Heiji reached overall criterion after nine sessions (864 trials). Although he learned to discriminate the second pair of primate faces before the second pair of human faces (92% correct for nonhuman primates in session 2; 92% for humans in session 7),
Figure 8. Acquisition of the first pairs of stimuli by Kiki. Average percent of face stimuli correctly selected per session for the first pair of human and nonhuman primate stimuli. They grayscale image of one pair of human and nonhuman primate faces were repeatedly presented in each session. Sessions consisted of 96 trials (48 trials of the human pair; 48 trials for the nonhuman primate pair). For each trial, the subject was required to select the image of the intact face presented on a touch screen. Criterion was reached in 37 sessions.

Figure 9. Acquisition of the second pairs of human and primate pictures and maintenance of performance by Kiki for the first pairs. Average percent of face stimuli correctly selected per session for the first and second pair of human and nonhuman primate stimuli. They grayscale image of the pairs were repeatedly presented in each session. Sessions consisted of 96 trials (27 trials for the first primate; 4 trials for the second primate pair; 21 trials for the first human pair; 24 times for the second human).
he was unable to maintain high performance, only achieving it once more in session 9. However, he maintained his performance for the second human pair. His maintenance performance was 87±8% for the first pair of nonhuman primate faces, and 87±12% for the first pair of human faces.

After reaching criterion performance for all pairs (both first and second), the third scrambled/intact pair of human and non-human primate pictures were introduced to Heiji, who reached overall criterion after 9 sessions. Although he learned to discriminate the second pair of primate faces before the second pair of human faces (92% correct for primates in Session 2; 92% correct for humans in session 7), he was unable to maintain high level of performance, only achieving it once more in session 9. However, he maintained his performance for the second human pair. His maintenance performance was 87±8% for the first pair of primate faces, and 87±12% for the first part of human faces.

Heiji was immediately able to discriminate the third scrambled primate face from intact primate face (92% in session 1) but did not reach criterion for discrimination in the third pair of human pictures (up to 83% in session 3). He alone has been tested with the third, fourth, fifth and sixth pairs. For the pairs introduced fourth, Heiji was immediately able to discriminate the scrambled human face from intact human face (96% in Session 1) but did not reach criterion for discrimination in the fourth pair of primate pictures (up to 88% in Session 4).

For the fifth pairs, he was able to reach criterion for both monkeys and humans in session 2. However, he did not maintain performance at criterion until the last several sessions, and appeared to be unable to discriminate the new monkey pair as well as the new human pair. He did
Figure 10. Acquisition of the first pairs of stimuli by Heiji. Average percent of face stimuli correctly selected per session for the first pair of human and nonhuman primate stimuli. They grayscale image of one pair of human and nonhuman primate faces were repeatedly presented in each session. Sessions consisted of 96 trials (48 trials of the human pair; 48 trials for the nonhuman primate pair). For each trial, the subject was required to select the image of the intact face presented on a touch screen. Overall criterion was reached in 12 sessions.

Table 7.

Maintenance of performance by Heiji for pairs after their initial introduction

<table>
<thead>
<tr>
<th></th>
<th>Human Pair</th>
<th>Nonhuman Primate Pair</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Percent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct ±SD (%)</td>
<td>First</td>
<td>Second</td>
</tr>
<tr>
<td></td>
<td>92±11</td>
<td>84±16</td>
</tr>
<tr>
<td></td>
<td>80±19</td>
<td>90±18</td>
</tr>
<tr>
<td></td>
<td>80±19</td>
<td>93±9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95±9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>88±15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>83±15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>87±9</td>
</tr>
</tbody>
</table>