ORIGINAL PAPER

Multiple perceptual strategies used by macaque monkeys for face recognition

**Katalin M. Gothard · Kelly N. Brooks ·
Mary A. Peterson**

Abstract Successful integration of individuals in macaque societies suggests that monkeys use fast and efficient perceptual mechanisms to discriminate between conspecifics. Humans and great apes use primarily holistic and configural, but also feature-based, processing for face recognition. The relative contribution of these processes to face recognition in monkeys is not known. We measured face recognition in three monkeys performing a visual paired comparison task. Monkey and humans faces were (1) axially rotated, (2) inverted, (3) high-pass filtered, and (4) low-pass filtered to isolate different face processing strategies. The amount of time spent looking at the eyes, mouth, and other facial features was compared across monkey and human faces for each type of stimulus manipulation. For all monkeys, face recognition, expressed as novelty preference, was intact for monkey faces that were axially rotated or spatially filtered and was supported in general by preferential looking at the eyes, but was impaired for inverted faces in two of the three monkeys. Axially rotated, upright human faces with a full range of spatial frequencies were also recognized, however, the distribution of time spent exploring each facial feature was significantly different compared to monkey faces. No novelty preference, and hence no inferred recognition, was observed for inverted or low-pass filtered human faces.

High-pass filtered human faces were recognized, however, the looking pattern on facial features deviated from the pattern observed for monkey faces. Taken together these results indicate large differences in recognition success and in perceptual strategies used by monkeys to recognize humans versus conspecifics. Monkeys use both second-order configural and feature-based processing to recognize the faces of conspecifics, but they use primarily feature-based strategies to recognize human faces.

Keywords Face recognition · Face perception · Non-human primates · Inversion effect · Configural processing · Feature-based processing · Scanpaths

Introduction

Survival in hierarchical primate societies requires individuals to identify members of their group quickly and accurately, interpret their facial signals, and respond to them with appropriate behaviors. Given its ethological value, face recognition in monkeys is likely supported by fast and efficient perceptual mechanisms. Three basic strategies have been proposed for face recognition. Individuals can be recognized using *holistic processing* by which faces are perceived as a whole, and each feature is processed in conjunction with all the other features (Maurer et al. 2002). Conversely, *feature-based processing* involves recognizing individuals by characteristic features alone. A third strategy, called *second-order configural processing*, uses the distances between features as well as the features themselves to discriminate between individuals who have similar features but different spacing between them (Diamond and Carey 1986). Holistic and second-order configural

K. M. Gothard (✉) · K. N. Brooks
Department of Physiology,
College of Medicine, The University of Arizona,
1501 N. Campbell Ave, Room 4104, Tucson,
AZ 85724, USA
e-mail: kgothard@email.arizona.edu

M. A. Peterson
Department of Psychology, The University of Arizona,
Tucson, AZ, USA

⚫ Springer

processing require familiarity and expertise, acquired developmentally (de Haan et al. 2002; Mondloch et al. 2006; Nelson 1995; Pascalis et al. 2002, 2005) or experimentally (Hills and Lewis 2006). Humans, chimpanzees, and monkeys have been shown to use holistic and/or configural processing (Parr and Heintz 2006; Parr et al. 2006; Dahl et al. 2007), but in macaques it is unclear whether face recognition is entirely holistic and/or configural, or whether it combines elements of configural and feature-based processing.

Standard stimulus manipulations can isolate different face recognition strategies. For instance, some authors suggested that inverting a face impairs holistic and configural processing more than feature-based processing (Yin 1969; Valentine and Bruce 1988; Maurer et al. 2002), while others showed that the inversion effect might be due to an increased difficulty to extract information from inverted faces, rather than a complete failure of configural/holistic strategy (Nachson and Shechory 2002; Rakover 2002; Sekuler et al. 2004). Spatial filtering can serve as a complementary manipulation that further discriminates between perceptual strategies. Removal of high-spatial frequency components impairs feature-based processing and some authors suggested that removal of low-spatial frequencies impairs holistic and configural processing (Goffaux et al. 2005; Sergeant 1986).

These manipulations alone do not always provide clear answers. Inversion, for example, has been shown by some experimenters to cause face recognition deficits in monkeys (Rosenfeld and Van Hoesen 1979; Overman and Doty 1982; Swartz 1983; Keating and Keating 1993; Phelps and Roberts 1994; Tomonaga 1994), while others found no impairment (Bruce 1982; Dittrich 1990; Gothard et al. 2004). The clarity of conclusions emerging from the face recognition literature in monkeys is further complicated by the diversity of tasks and stimuli used. A matching-to-sample task, for example, often requires over 1,000 training trials (Parr et al. 2000), hence overtraining effects and reward-seeking strategies might complicate the interpretation of the results. In contrast, the visual paired comparison task (VPC) requires minimal or no training (Gunderson and Swartz 1985; Pascalis and Bachevalier 1999; Gothard et al. 2004). Whatever the task, however, it is essential to use as sample and match different views of the same face. The recognition of individuals in the natural environment most likely requires view-invariant representation. Using the same picture reduces the perceptual demands of the task and is less informative about perceptual mechanism (Lacreuse and Herndon 2003; Gothard et al., 2004). Our goal was to understand the strategy or combination of strategies monkeys use to recognize conspecific and human faces. Given that different stimulus manipulations favor or impair different face recognition strategies, we tested each

monkey with a series of axially rotated, inverted, or spatially filtered faces in order to answer three specific questions:

1. Do rhesus monkeys use view-invariant representations for face recognition? If they do, recognition should be intact over axial rotation.

2. Do rhesus monkeys use holistic, second-order configural, and/or feature-based face processing strategies? If monkeys use holistic/configural processing face recognition is expected to be impaired by inverting the stimuli *and* by removing low-spatial frequencies from images (high-pass filtered images). Alternatively, if monkeys use feature-based processing, recognition should be impaired for low-pass filtered faces, where the details of features are lost, but intact for high-pass filtered faces.

3. Do monkeys use different visual strategies for humans versus conspecifics? Holistic and second-order configural processing, which depend on expertise, are expected to favor conspecifics (Dufour et al. 2006; Pascalis and Bachevalier 1999; Wright and Roberts 1969). If they do, we should observe an inversion effect for faces of monkeys but not humans, and better recognition of low-pass filtered monkey faces than human faces. If monkeys use feature-based processing for human faces but not monkey faces, we expect better recognition for high-pass filtered human faces than monkey faces.

We used the VPC task to quantify face recognition because this task does not require training, rather, it exploits the monkeys' spontaneous preference for novelty (Fantz 1964; Fagan 1972). The VPC task relies on the premise that the animal retains a representation of a face that can be compared later to a different image of the same face or to a new face. Successful recognition of the familiar face causes spontaneous orientation toward the novel face, termed a "novelty preference" (Nelson 1995; Overman et al. 1992; Richmond et al. 2007). As a fine-grained indication of recognition strategy we quantified the proportion of time monkeys spent looking at particular face regions (e.g., Eyes, Mouth, and "Other" regions, see below) during recognition, and examined whether a novelty preference was obtained for individual regions.

Finally, we examined whether the monkeys spent more time looking at the eye regions of the match faces, regardless of whether or not they showed a novelty preference. When viewing faces, human subjects look preferentially at the eyes (Vinette et al. 2004), and ERP components typical of face perception are observed only when eyes are present in the face images (Bentin et al. 2002; Schyns et al. 2002). Therefore, we reasoned that if the monkeys perceive images as faces, regardless of whether or not they recognize the individuals, they would look preferentially at the eyes.

Methods

Subjects

Three 6–8 years old (8–13 kg), male rhesus monkeys (*Macaca mulatta*), T, Q, and H, served as subjects. The monkeys were naïve to the visual paired comparison (VPC) task and were unfamiliar with the monkey and human faces used as stimuli. All three animals were born and mother-raised in outdoor enclosures at the California National Primate Research Center (Davis, CA, USA). They have acquired, therefore, extensive exposure and social experience with the faces of their conspecifics and to a lesser extent with human faces. During testing they were housed individually in a rhesus monkey colony room with auditory and some visual contact with one another. All experimental procedures were performed in compliance with the guidelines of the National Institutes of Health for the use of primates in research and were approved by the Institutional Animal Care and Use Committee at the University of Arizona. In preparation for recording eye movements, each monkey was fitted with a head fixation device attached to the skull under isoflurane anesthesia.

Behavioral task and training

The monkeys were exposed for 3 s to a pair of images known as *the familiarization pair*, which comprised two images of the same face (monkey or human), one on each side of center. The faces were oriented so that they both looked toward the center of the screen. The individuals in the photographs were shown with an averted gaze. After a delay of 1 s, during which time no visual stimuli were shown, a second pair of images was presented—*the test pair*. The test pair was displayed for 3 s and comprised an image of the individual from the previous familiarization pair and a novel individual (Fig. 1). Again, both individuals were oriented toward the center of the screen, and were shown with gaze averted. There were equal number of males and females and two of the males were bearded. The two individuals in the test pair resembled each other and matched as closely as possible in terms of facial hair, color and head size. The humans wore a surgical cap to eliminate hair cues. Scanpath analyses determined the target of the first saccade on the test pairs, the dwell time on the target of the first fixation, and also the total looking time spent exploring the *Eyes*, *Mouth*, and *Other* facial features. We examined the mean durations monkeys looked at each of these regions for both the novel face and the familiar face on each trial. Recognition of the familiar face is inferred from longer looking times at one or all of the face regions in novel compared to familiar faces (Fagan 1972). The quantification of novelty preference for sub-regions of the
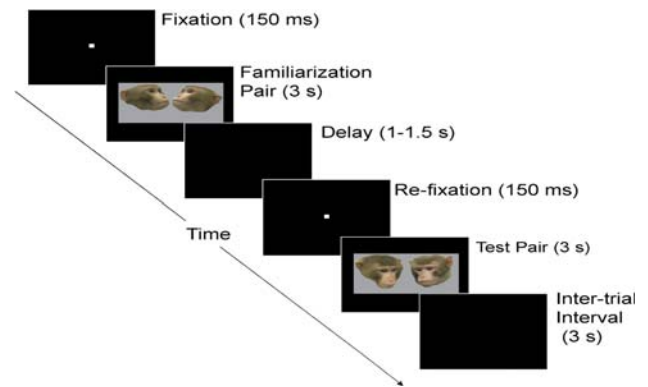

Fig. 1 Detailed task protocol: Each trial started with the central display of a white fixation icon, subtending 0.25° visual angle. Subjects were required to fixate at this icon within 1 s of its appearance and maintain fixation for at least 150 ms (fixation). If they failed to do this, the icon was removed from the monitor and a time-out period of 2 s was imposed. Immediately following a successful fixation the familiarization pair was displayed for 3 s (familiarization). The monkeys were free to explore the two images of the familiarization pair, but to obtain a drop of fruit juice, they were required to maintain their gaze inside the boundary of the image pair (otherwise the monitor went blank for 2 s and no juice reward was delivered). Following a delay of 1 s, plus a random delay of up to 500 ms, the fixation icon reappeared (re-fixation). If the monkeys failed to fixate on the spot within 1 s, the trial was aborted, and a 2 s time-out period was imposed. If, however, fixation occurred within the required period, the second pair of images, the test pair, was displayed for 3 s (test). The monkeys were required to maintain gaze within the boundary of the image pair for the entire duration of the display. At the successful completion of the test phase, the monkey received two drops of juice reward. Juice reward was always paired with an auditory stimulus, indicating a correct trial. After a 3 s inter-trial interval, a new trial began with the display of the fixation spot. Error trials were repeated later in an experimental session

face allows the discrimination between possible strategies used by the monkeys to recognize faces.

Behavioral training was achieved in three steps: (1) fixation training, (2) gaze maintenance training, and (3) viewing pairs of images displayed in the VPC task format. During training and data collection, a subject was seated in a primate chair in front of a 17-inch computer monitor placed 59 cm from the monkey's eyes. At this distance, 1 cm on the monitor corresponds to 1° of visual angle (dva). The chair and monitor were enclosed in a booth. For correct responses, monkeys were rewarded with a drop of fruit juice mixed with applesauce. Incorrect responses were followed by a 2 s time-out period, during which the screen remained blank and no reward was given.

1. Fixation training. Subjects were first trained to orient to and fixate for 200 ms on small white squares that subtended 0.25 dva in the center of the computer monitor.
2. Gaze maintenance training. The monkeys scanned a 24 × 12 dva image without looking outside its boundaries. The duration of the display was increased gradually from 1 to 3 s. Images were preceded by the fixation icon. Trials were separated by a 3 s inter-trial interval.

3. VPC training. The monkeys were trained to look at two image pairs separated by a 1-s delay. Image pairs consisted of two adjacent 12 × 12 dva images that depicted dogs, cats, or complex scenes. Performance criterion was met when the monkeys fixated for at least 200 ms on fixation icons and maintained gaze within the boundaries of training images on 90% of the trials. All three monkeys learned this task within 4 weeks of daily training. A schematic of the task is shown in Fig. 1.

Experiments

Experiment 1 was designed to test whether rhesus monkeys use view-invariant representations for face recognition. The stimuli consisted of monkey faces and human faces in three-fourth views with full color and spatial frequency components. The "*familiarization pair*" contained two different three-quarter views of the same human or monkey; the two faces were oriented toward the midline. The "*test pair*" contained a third three-quarter view of the same individual and a three-fourth view of a different individual. This was the case for all pairs of faces for all the experiments. An additional difference between the faces in the familiarization and test pairs was the upward or downward gaze direction of the model (Fig. 2).

Experiment 2 tested for an inversion effect for human or monkey faces. We used gray-scale monkey and human faces with full spatial frequency components. The familiarization pair contained two upright three-quarter views of an individual; unlike Experiment 1, these two views were mirror images of each other. The test pair contained an inverted (i.e., upside down) copy of one of the images
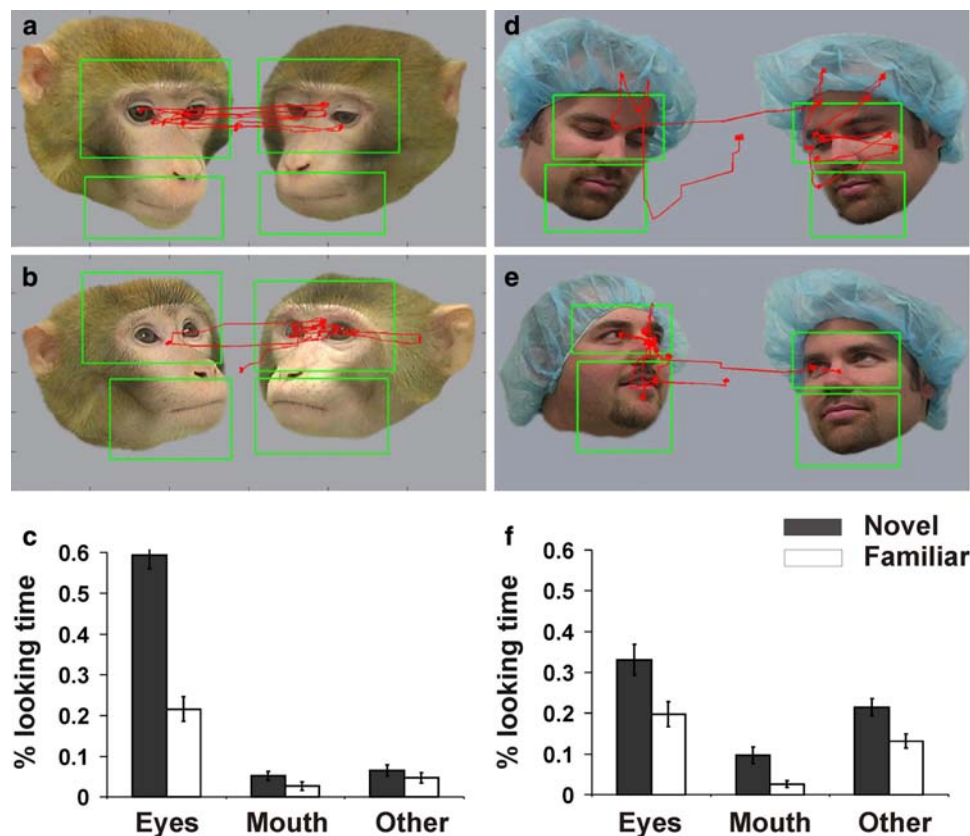

Fig. 2 Face recognition based on view-invariant representation. **a** Example scanpath (*in red*) on a familiarization pair of monkey faces. The *green rectangles* represent the boundaries of eye and mouth regions of interest. **b** Scanpath on the corresponding test pair. The image on the left is a third view of the monkey shown in the familiarization pair (*above*). Note that in addition to head orientation, gaze direction is also different between the familiarization and test images. The scanpath on the familiarization consisted of a series of saccades and fixations directed at the eyes of the two monkeys; 64% of the time the viewer monkey explored the face on the left. On the test pair, the first saccade was directed toward the eyes of the novel monkey (*right side*); the second saccade was aimed to the eyes of the familiar face; the monkey then returned to explore the *eye* and *other* regions of the novel face for the rest of the allowed viewing time. **c** Average (± SE) looking times for 20 test pairs of monkey faces indicating the proportion of total time spent looking at the each face region (Eyes, Mouth, and Other) on novel and familiar monkey faces. **d** Example scanpath on a familiarization pair of human faces. **e** Example scanpath on the corresponding test pair of human faces. The face on the left is the novel individual. Regions of interest on both familiarization and test pairs are indicated by green rectangles. **f** Average (±SE) looking times for 20 test pairs of human faces indicating the fraction of total looking time spent viewing each face region (Eyes, Mouth, and Other) on human faces

from the familiarization pair adjacent to an inverted novel individual (Fig. 3).

Experiment 3 tested whether monkeys can recognize human and monkey faces using high-spatial frequency components alone. Feature-based recognition should be intact in high-spatial frequency images of faces, whereas, second-order configural processing should be impaired when the low-spatial frequency component of the face are missing from the image. In this experiment, the familiarization images were the same as those used in Experiment 2. The test images were high-pass filtered with a filter radius of 2.0 pixels (Fig. 4) using Photoshop 7.0. This manipulation makes the faces look like sharpened line drawings. The brightness and contrast of the familiarization and test images was matched using Gimp 2.0.

Experiment 4 tested whether monkeys can recognize human and monkey faces using low-spatial frequency information alone. Second-order configural processing should be intact for low-pass filtered faces, whereas feature-based processing, which depends on high frequency components, should be impaired. The same familiarization pairs were used as in Experiments 2 and 3. The test pairs consisted of a copy of the previously seen face shown side by side with a novel image; a low-pass spatial frequency filter was applied to both images shown in the test pairs (Fig. 5). We used a Gaussian blur with a radius of 15.0 pixels (Adobe Photoshop 7.0) to remove detailed information of facial features (carried by high-spatial frequencies). The low-pass filtered images looked blurred (Fig. 4).
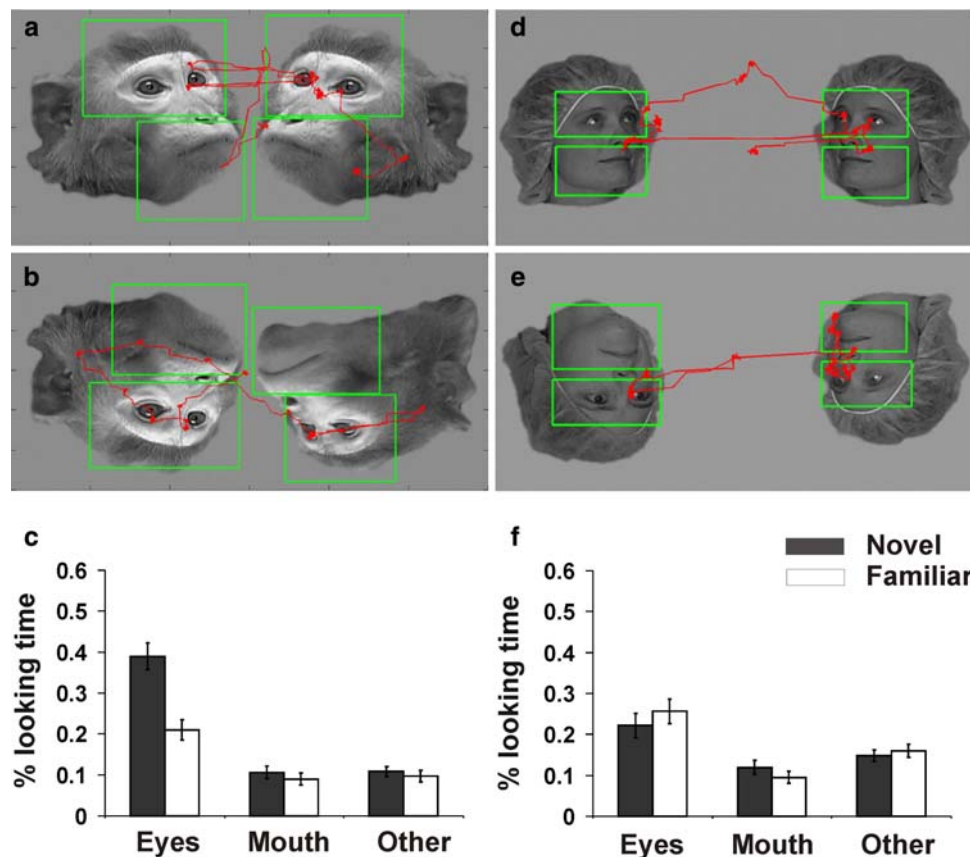

Fig. 3 Recognition of inverted monkey and human faces. **a** Example scanpath (*in red*) on a familiarization pair of upright monkey faces (mirror image views of the same monkey). The *green rectangles* represent the boundaries of regions of interest. The looking time on the two upright monkey faces was evenly distributed (51% on left and 49% on right). **b** Scanpath on the corresponding test pair. The image on the left is an inverted copy of the face in the familiarization pair while the image on the right is novel. For this particular trial, the first saccade on the test pair was directed at the eyes of the familiar face, followed by several small saccades to various features of the familiar face that added up to 84% of the looking time. The viewer monkey crossed over to the novel face only during the last 500 ms of the scan. **c** Average (±SE) looking times for 20 test pairs of monkey faces indicating the proportion of total time spent looking at the each face region (Eyes, Mouth, and Other) on novel and familiar monkey faces. **d** Example of scanpath on a familiarization pair of upright human faces, depicting mirror images of the same individual. **e** Scanpath on the corresponding test pair of inverted human faces. The face on the left is the novel individual. **f** Average (±SE) looking times for 20 test pairs of human faces indicating the fraction of total looking time spent viewing each face region (Eyes, Mouth, and Other) on human faces
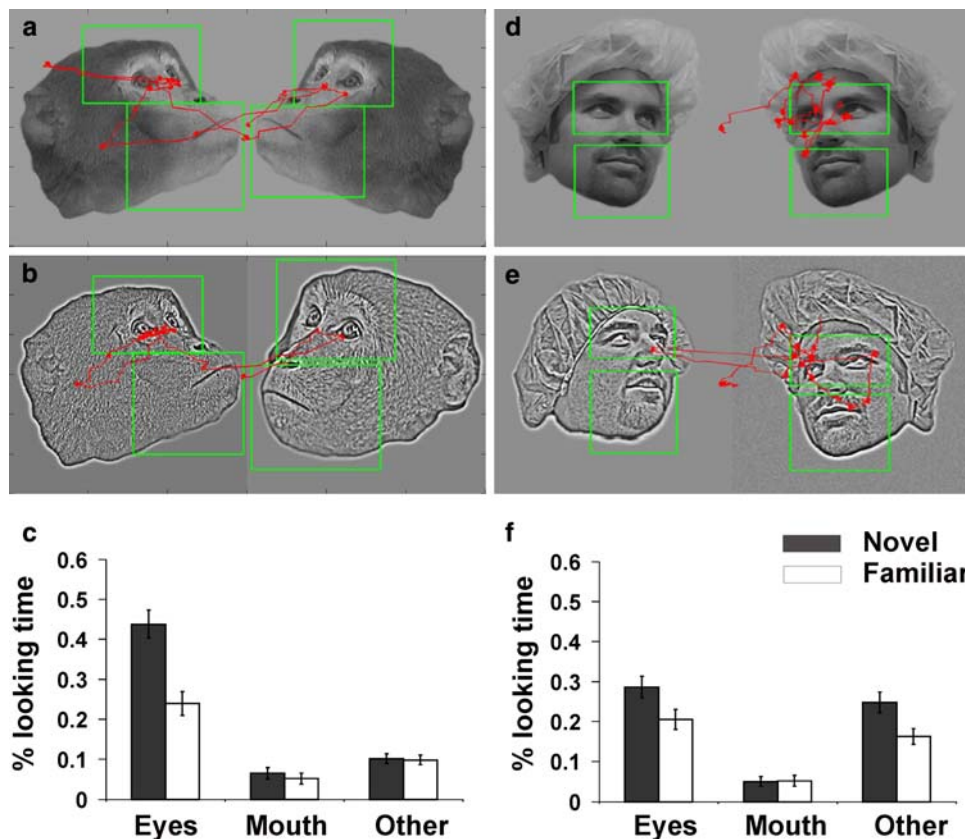
Fig. 4 Recognition of high-pass filtered monkey and human faces. **a** Example scanpath (*in red*) on a familiarization pair of full frequency, mirror-view images of the same monkey. The *green rectangles* represent the boundaries of regions of interest. **b** Scanpath on the corresponding test pair of high-pass filtered monkey faces. The image on the right is a high-pass filtered copy of the familiar face, while the image on the left contains a novel face. **c** Average (±SE) looking times for 20 test pairs of monkey faces indicating the proportion of total time spent looking at the each face region (Eyes, Mouth, and Other) on novel and familiar monkey faces. **d** Example scanpath on a familiarization pair of full-frequency, mirror-views of the same human face. **e** Scanpath on the corresponding test pair of high-pass filtered human faces. The image on the left is high-pass filtered copy of the familiar face, while the image on the right is a high-pass filtered novel face. **f** Average (±SE) looking times for 20 test pairs of human faces indicating the fraction of total looking time spent viewing each face region (Eyes, Mouth, and Other) on human faces

Low- and high-pass filtered images were equated within and across sets for brightness and contrast using image-editing programs (Adobe Photoshop 7.0 and Gimp 2.0).

The monkey faces used as stimuli (20 unfamiliar monkeys) were selected from an extended library of digitized monkey faces (Gothard et al. 2004). All external cues were removed from the face stimuli using Adobe Photoshop 7.0. Images of 20 unfamiliar humans were extracted from video footage recorded with a digital camera (Canon GL1) on Sony mini premium digital cassettes set to maximum resolution. All stimuli were digitized images that subtended 12×12 dva. The human faces were shown with blue surgical head covers, to prevent the monkeys from using hairstyle as a cue for face recognition. Due to the social significance of direct gaze in monkeys, all faces were shown in three-fourth profiles with either upward or downward gaze. All faces were shown on a gray matte background.

For all experiments human faces and monkey faces were tested in separate blocks. A block consisted of 20 trials with 10 sets of images. Each trial (familiarization pair, delay, and test pair) was performed twice, with the novel face on the left and the right side. Balancing the presentation of the novel image on the left and right eliminated the effect of possible left–right biases. The sequence of trials within a session was randomized. Only successful trials were used for data analysis. All three monkeys completed all 20 trials in each experiment with one exception: one monkey was not tested with the monkey faces in Experiment 1. Because the results from the other two monkeys were consistent, we do not believe this omission limits the generality of the results of Experiment 1. In VPC task, absolute familiarity with the stimuli has not been shown to influence the outcome; we tested, therefore, each monkey 2–3 times with each stimulus set and analyzed the data from the experiments where monkeys made the fewest errors of initiating a
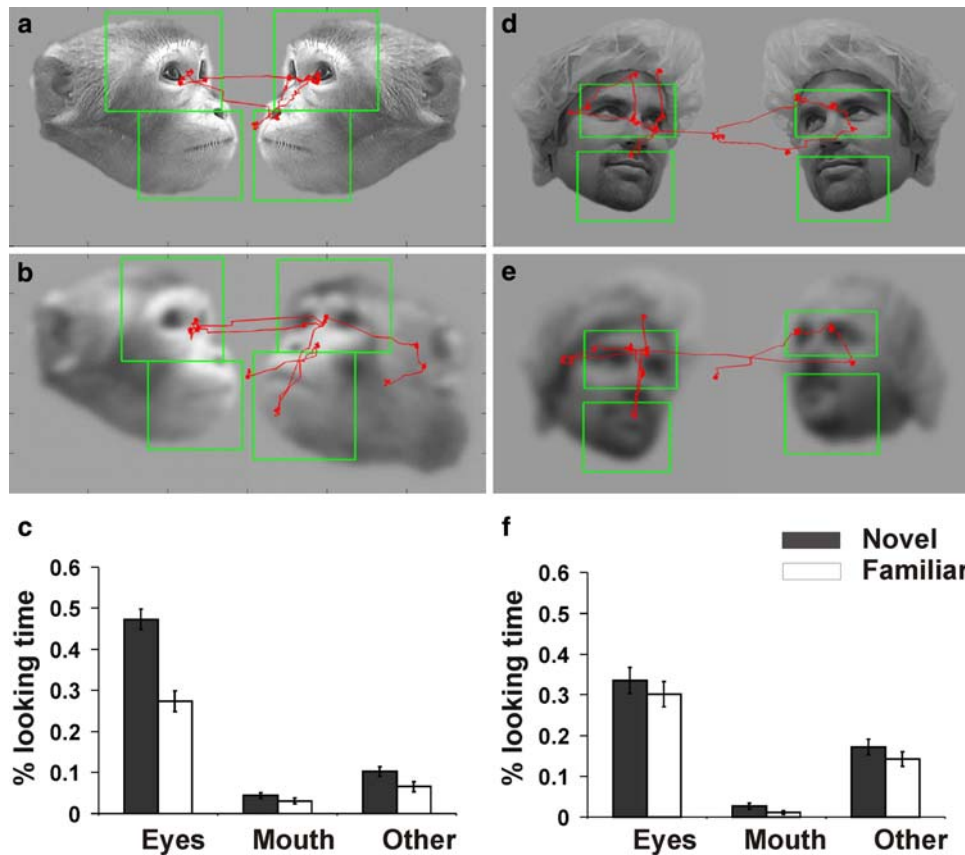
Fig. 5 Recognition of low-pass filtered monkey and human faces. **a** Example scanpath (*in red*) on a familiarization pair of full frequency, mirror-view images of the same monkey. The *green rectangles* represent the boundaries of regions of interest. **b** Scanpath on the corresponding test pair of low-pass filtered monkey faces. The image on the left is a low-pass filtered copy of the familiar face, while the image on the right is a novel face. **c** Average (±SE) looking times for 20 test pairs of monkey faces indicating the proportion of total time spent looking at the each face region (Eyes, Mouth, and Other) on novel and familiar monkey faces. **d** Example scanpath on a familiarization pair of full-frequency, mirror-views of the same human face. **e** Scanpaths on the corresponding test pair of low-pass filtered human faces. The image on the left is low-pass filtered copy of the familiar face, while the image on the right is a low-pass filtered novel face. **f** Average (±SE) looking times for 20 test pairs of human faces indicating the fraction of total looking time spent viewing each face region (Eyes, Mouth, and Other) on human faces

trial (fixation) or terminating the trial before the required looking time has expired (looked outside the boundary of images).

Measurement of eye movements

Eye movements were tracked with a resolution of 0.25 dva using an infrared eye tracker (ISCAN, Inc., Burlington, MA, USA), and digitized at 500 Hz using the CORTEX data acquisition program (NIMH-supported freeware from the website: http://www.cortex.salk.edu). Scanpaths are defined here as digitized eye movements calibrated in degrees of visual angle and superimposed on stimulus images. To quantify time spent viewing various facial features, an experimenter, blind to the monkey's scanpaths, drew rectangles, delineating regions of interest, around facial features (examples of these rectangles are shown in Figs. 2, 3, and 4). Regions of interest included: (1) *Eyes*, (2) *Mouth*, and (3) *Other* (areas outside of the eye and mouth regions were, by default, classified as "*Other*"). The margins of these regions of interest slightly exceeded the boundaries of the named feature to ensure that a fixation on the margin of that feature would be included in viewing time spent on that feature. The rectangles around the eyes, on both human faces and monkey faces, were drawn to include the eyebrows as well as a region beneath the eyes. Due to the relatively great variability of horizontal and vertical rotation of the faces, in some cases, portions of the nose may have fallen into the rectangle surrounding the eyes; however, we rarely noticed subjects looking at the nose and did not include it as a separate region of interest. For all human faces and monkey faces, the *Mouth* region always included the area just below the nose and extended just beyond the chin. Due to variability in the size of facial features across human faces and monkey faces, the regions of interest varied in size across stimuli. After overlaying a

scanpath on a stimulus image, the looking time was quantified using MATLAB (Math Works, Inc., Natick, MA, USA) custom designed programs that added all the points that fell within the boundaries of the regions of interest.

Data analysis

Only trials on which the monkeys maintained their gaze for 3 s within the boundary of the stimulus images were considered successful. The percent looking time at each face region on the novel and familiar half of the image was averaged across 20 successful trials (Recall that 10 sets of faces were presented two times each, with the novel face once on the right and once on the left). For each test pair we assessed the duration of time each monkey looked at each face (i.e., the familiar face and the novel face). These durations summed up to less than 3 s because the eyes remained fixated at the location of the previously displayed fixation icon for 150–250 ms at the beginning of each trial. Looking time was therefore quantified as % looking time at each of the three face regions (*Eyes*, *Mouth*, and *Other)* for each of the two faces, novel and familiar. Although the overall % times looking at novel versus familiar faces is necessarily complementary, the % time looking at individual features of novel versus familiar faces are not. We examined whether novelty preference depends on face region, via a 4-way repeated measures ANOVA where the factors were Subjects (3 levels, H, T, Q), Species (2 levels, monkeys and humans), Face regions (3 levels, *Eyes*, *Mouth*, *Other*), and Novelty (2 levels, novel and familiar).

Results

Experiment 1: Recognition of faces based on view-invariant representations

For the upright faces we first performed an ANOVA comparing performance with monkey and human faces on the data obtained from the two monkeys (T and Q) that saw both monkey and human faces (H was not tested in the upright monkey face condition). We found a significant 3-way interaction among Species, Region, and Novelty, $F_{2,76} = 12.23$, $P < 0.001$: On monkey faces the Subjects looked significantly longer at the *Eyes* of novel faces (0.59 vs. 0.22, $P < 0.001$). They looked at the *Mouth* and the *Other* regions only rarely (mean = 0.04), and did not look longer at these regions in novel than familiar faces. In contrast, they showed a novelty preference for all three regions of human faces (*Eyes:* 0.33 vs. 0.20, *Mouth*: 0.10 vs. 0.03, and *Other:* 0.22 vs. 0.13). Note, however, that T and Q looked less often at the *Eyes* of human faces than monkey faces.

The ANOVA also revealed a 3-way interaction among Species, Region, and Subjects, $F_{2,76} = 5.6$, $P = 0.005$, indicating that whereas both Subjects looked primarily at the eyes of conspecifics, regardless of novelty, only Q showed this pattern for humans; T looked approximately equally often at the *Eyes* and *Other* regions. Thus, Q showed a face-specific pattern of looking at human faces as well as monkey faces, but T did not.

We conducted a second ANOVA on the data from all three Subjects examining performance with human faces only. An interaction between Region and Subject, $F_{4,114} = 6.76$, $P < 0.001$, indicated that like the other monkeys, H showed a novelty preference for all regions of human faces, and like Q (and unlike T), he looked primarily at the eyes of human faces.

This first experiment established that monkeys recognize conspecifics and human faces across different orientations, albeit using different facial features for recognition. This may indicate the use of different perceptual strategies for face recognition across species. It also revealed that T differs from H and Q in that for human faces, he does not show the face-specific pattern of looking at the *Eyes* more than the other two face regions. Finally, we found that monkeys looked at the *Eye* regions more often when they viewed monkey faces rather than human faces.

In the following experiments we investigate what type of processing underlies the viewpoint invariant recognition for each species. We begin by investigating whether a novelty preference is obtained for inverted monkey and human faces after familiarization with upright faces. The outcome of the inversion test might better inform about possible strategies.

Experiment 2: Recognition of inverted monkey and human faces

As in Experiment 1, we obtained a 3-way interaction among Species, Region, and Novelty $F_{2,112} = 7.44$, $P = 0.001$. Here the subjects showed a novelty preference for conspecific *Eye* regions only (0.39 vs. 0.21, $P < 0.001$); they showed no novelty preference for human faces. A marginal 4-way interaction among Species, Region, Novelty, and Subjects, $P = 0.063$ led us to perform a follow-up ANOVA examining the novelty preference for inverted monkey *Eye* regions. In this ANOVA an interaction between Novelty and Subject, $F_{2,56} = 4.17$, $P = 0.02$, revealed that only Q showed a statistically significant novelty preference for the conspecific *Eye* regions (Q: 0.57 vs. 0.18, $P < 0.001$; H; 0.36 vs. 0.28, $P > 0.05$; T: 0.23 vs. 0.18, $P > 0.05$) indicating that only Q recognized the inverted monkey faces.

The overall ANOVA also showed an interaction between Species and Region $F_{2,112} = 14.76$, $P < 0.001$: the subjects looked longer at the *Eye* regions than the *Mouth* or

the *Other* face regions, but the difference was larger for monkey than human faces. A significant 3-way interaction among Species, Region, and Subjects $F_{4,112} = 2.80$, $P = 0.029$, moderating this effect led us to examine whether all Subjects looked longer at the *Eye* region of human faces. We obtained a subjects-by-region interaction: $F_{4,112} = 29.93$, $P < 0.001$: both H and Q looked longer at the *Eye* region of human faces (i.e., they showed the face-specific looking pattern), whereas T looked longer at the *Other* region.

We conclude that inversion impaired recognition. Only one subject (Q) recognized inverted monkey faces and none of the subjects recognized inverted human faces. For the most part, subjects perceived the inverted stimuli as faces, as revealed by a tendency to look at the *Eyes* more than the *Mouth* or the *Other* face regions (only T failed to look longer at the *Eyes* of human faces).

Holistic and second-order configural processing are impaired by inversion; feature-based processing can also be impaired, albeit to a lesser extent. Hence, despite the inversion effects it remains possible that monkeys were using a feature-based strategy to recognize the upright stimuli. Also, given that Q can recognize inverted monkey faces, it remains possible that he is using a feature-based strategy exclusively. In the following experiment we examined the extent to which feature-based processing contributes to face recognition using high-pass filtered test stimuli from which the low spatial frequency components were removed. (Recall that low-spatial frequency components support second-order configural processing, whereas high spatial frequency components support feature-based processing).

Experiment 3: Recognition of high-pass filtered monkey and human faces

Once again, the ANOVA revealed a 3-way interaction among Species, Region and Novelty $F_{2,114} = 4.74$, $P = 0.01$. The Subjects showed a novelty preference for conspecifics by looking more often at the eyes of novel than familiar monkey faces. They also looked more frequently at the *Eyes* of conspecifics than the *Mouth* or *Other* regions, the looking pattern we take as face-specific. In contrast, the Subjects looked approximately equally often at the *Eyes* and *Other* regions of the human faces. They showed a novelty preference for human faces by looking more often at the *Other* regions of novel than familiar faces (0.25 vs. 0.16, $P = 0.046$); they did not show a novelty preference for the *Eyes* of high-pass filtered human faces.

A significant 3-way interaction among Subjects, Species, and Region $F_{4,114} = 8.29$, $P < 0.001$, indicated that Q looked at the eyes of high-pass filtered monkeys more than either H or T did; all three Subjects looked equally often at the eyes of humans (and not more often than at the *Other* regions).

Despite the presence of a novelty preference for both monkey and human faces, these results converge with those of Experiment 1 to suggest that individuals of the two species are recognized via different strategies. The subjects' ability to differentiate between individual high-pass filtered monkey faces is expressed by a novelty preference in the *Eye* region. In contrast, their ability to differentiate between individual high-pass filtered human faces is evident in a novelty preference in the *Other* face region. The results of Experiment 3 indicate that the monkeys use at least some feature-based processing to recognize individuals of both species. Even so, the subjects rely on a different set of features for conspecifics versus humans. In addition, the absence of a face-specific looking pattern for high-pass filtered human faces raises the possibility that the subjects did not perceive them as faces.

We turn now to the complementary experiment using low-pass filtered test images to examine whether the monkeys can also use second-order configural processing for monkey and human face recognition.

Experiment 4: Recognition of low-pass filtered monkey and human faces

As in all previous experiments we observed a 3-way interaction among Species, Region, and Novelty, $F_{2,114} = 5.77$, $P < 0.004$ (find exact P). Monkeys revealed a novelty preference for the *Eyes* of conspecifics only (0.47 vs. 0.27, $P < 0.001$) and not for *Mouth* or *Other* face regions and they failed to show a novelty preference for humans for any face region.

There was also 3-way interaction among species, region, and subjects, $F_{4,114} = 2.46$, $P < 0.049$: All Subjects looked longer at the *Eyes* of monkey faces than the *Mouth* or *Other* regions. Subjects H and Q showed the same pattern for human faces, whereas T's looking times at the *Eyes* and *Other* regions were not significantly different ($P > 0.07$). T's failure to look less often than the other two subjects at the *Eye* regions was evident in previous experiments as well.

Low-pass filtering disrupts feature-based processing and is purported to preserve second-order configural processing. All monkeys had preserved recognition for low-pass filtered monkey faces, suggesting that they can use second-order configural processing for conspecifics. Interestingly, they failed to recognize low-pass filtered human faces, indicating that second-order configural processing alone was not sufficient to support recognition of human faces.

Discussion

This series of experiments show that monkeys can individuate faces of conspecifics and humans, but they do so using

different, but not mutually exclusive perceptual mechanisms. Specifically, they use both second-order configural and feature-based processing to recognize the faces of individual conspecifics, but they use primarily feature-based strategies to recognize human faces.

In Experiment 1, all three monkeys showed viewpoint-invariant recognition of upright human and monkey faces. A 3-s exposure to two 3/4 views of an unfamiliar face provided sufficient information about the geometry of the features to allow the monkeys to later identify the individual from a third viewpoint. The phenomenon of viewpoint-invariant recognition was described by Marr (1980) and later by O'Toole et al. (1999), predicted by theoretical models (Troje and Bulthoff, 1998), and confirmed in humans by functional imaging (Pourtois et al., 2005). By analyzing the specific face regions at which monkeys looked, we found evidence that they employ different strategies to recognize conspecifics and humans. Specifically, conspecific recognition was evident in longer looking times at the *Eyes* of the novel monkey rather than the familiar monkey, whereas human recognition was evident in longer looking times in all of the face regions of the novel faces compared to the familiar faces. Additional experiments were conducted to attempt to determine the nature of those different strategies.

In Experiment 2, two of the three monkeys failed to show a novelty preference for inverted faces, regardless of species; one monkey, Q, showed a novelty preference for inverted monkey faces, but not for inverted human faces. Thus, the recognition strategies used by the monkeys are disrupted by inversion, but taken alone, this inversion effect does not allow us to reach conclusions regarding which recognition strategy they used. Inversion can impair holistic processing, and using composite faces, Dahl et al. (2007), found strong signs of holistic face processing in monkeys. Inversion can also disrupt second-order configural processing (Leder and Bruce 2000; Rhodes et al. 2006), perhaps more than any other perceptual strategy. Finally, inversion can disrupt feature processing as well (Maurer et al. 2002). Whatever type—or types—of processing are impaired in monkeys by inversion, the ability to perceive the test stimuli as faces was intact in all monkeys for conspecific faces and in two of the three monkeys for human faces. Finally, the fact that one monkey, Q, remained able to discriminate between inverted novel and familiar monkey faces, raises the possibility that he may either have more experience than the other monkeys with inverted faces, or use a strategy that better survives inversion the others. One candidate is feature-based processing as it may be less impaired by inversion than the other strategies.

Feature-based recognition strategies were tested with high-pass filtered faces. All monkeys showed a novelty preference for monkey faces by looking more at the *Eye*

regions of novel than familiar monkeys. They also showed a face specific pattern by looking predominantly at the *Eye* regions of monkey faces. In contrast, although the monkeys did show a novelty preference for high-pass filtered human faces, it was evident in differential looking times at the *Other* face region of novel versus familiar human faces rather than the *Eye* region. Furthermore, rather than looking longer at the Eye region of the human faces overall, they looked approximately equally often at the *Eye* and *Other* regions of human faces. That the monkeys show a novelty preference for high-pass filtered faces of both species suggests that they can use features to recognize individuals. The feature-based strategies used for conspecifics and humans rely, however, on different features.

Second-order configural processing strategies were tested with low-pass filtered images. All monkeys showed a novelty preference for faces of conspecifics, but none showed a novelty preference for human faces in this condition. Thus, these monkeys can use configural processing for faces of individuals of their own species, but not for human faces. Despite their inability to recognize individual humans, at least two of the monkeys (appear to perceive the low-pass filtered human faces as faces, evidenced by looking longer at the eyes than the other face regions.

A collective interpretation of the result reported here, and the novel aspect of this study is that monkeys use a combination of strategies to recognize conspecific faces; primarily holistic/configural and, to a lesser extent, feature-based processing for monkeys and a different, more limited, combination of strategies for humans (Dufour et al. 2006; Parr et al. 1999; Tomonaga 1994). In terms of the specific perceptual strategies used for human faces, our results suggest two possibilities: (1) Monkeys use both holistic and feature-based but not second-order configural mechanism for human faces. This is supported primarily by their failure to recognize low-pass filtered human faces. (2) Monkeys use only feature-based strategies for human face, and this strategy is not robust enough to survive inversion. Further tests are required to discriminate between these, or other possibilities. The mixed strategies use by monkeys and humans for face recognition might derive from the observation that faces, whether upright, inverted, axially rotated, or spatially filtered, engage the same or partially overlapping neural substrates. For example, human fMRI studies show that the fusiform area in humans is engaged in processing both upright and inverted faces (Kanwisher et al. 1998) as well as other stimuli requiring expertise-level mechanisms (Gauthier et al. 2003; Tarr and Gauthier 2000). A widely distributed network of face-responsive neurons in the visual areas of the temporal lobe support view-invariant face recognition and extract identity from degraded images (Wallis and Rolls 1997). The use of configural processes for conspecific faces but not human faces may occur because

monkeys have more extensive experience with conspecifics than with humans (REFS).

These data are consistent with the view that the various perceptual strategies are not mutually exclusive; rather more than one strategy might be at play depending on stimulus manipulation (Ghazanfar and Santos 2004) and species. The idea that monkeys can use both holistic/configural and feature-based processing has been demonstrated by Vermeire and Hamilton (1998), who found a right-hemispheric advantage for holistic/configural face processing and left-hemispheric bias for feature-based processing in split-brain rhesus monkeys. Humans also use a combination of mutually non-exclusive strategies to identify faces (Yovel and Kanwisher 2004). Sekuler et al. (2004) concluded that discriminative features (i.e., the eyes) are not used differentially to recognize upright compared to inverted faces; they concluded that, "the primary difference between processing upright and inverted faces is quantitative rather than qualitative." This proposal could be directly confirmed in monkeys by using tasks where the subjects are allowed to explore the faces as long as necessary to make a correct operation response.

Successful recognition and use of facial information has been frequently associated with significant biases for exploring certain facial regions, most commonly the areas of the eyes (Bentin et al. 2002; Ghazanfar and Santos 2004; Vinette et al. 2004; Schyns et al. 2002). A tendency to look at the eyes is present even when monkeys look at videos of vocalizing conspecifics (Ghazanfar et al. 2006). In our experiments, on the majority of trials and across all stimulus manipulations, the monkeys spent the largest percentage of time exploring the *Eye region* of conspecifics, and in Experiments 1 and 3 (upright and low-pass test stimuli), they explored the Eyes of humans more than the Mouth and Other face regions. This finding replicates earlier similar results (Kyes and Candland 1987; Nahm et al. 1997; Emery 2000; Gothard et al. 2004; Guo et al. 2006; Dahl et al. 2007; Guo 2007). Furthermore, it suggests that the monkeys recognized these images as faces, and attended to their key features. Feature preference, i.e., the percentage of time spent in exploring various facial features was not exclusive for conspecifics but was less prominent for human faces.

There were individual differences between the monkeys in the degree to which they explored the *Eye* region, both of conspecifics and of humans. Although we have described the tendency to look predominantly at the *Eye* region, as an index of whether or not the stimuli were perceived as faces, differences may also arise because of the hierarchical nature of macaque societies. Although all monkeys and humans used as stimuli were unfamiliar to the subject monkeys, it is likely that the viewer monkeys automatically estimated the status relationship between the stimulus and self. A high-ranking, confident individual, such as monkey H, might

perceive himself in an advantageous position when confronted with the face of an adult male or a human, but an anxious monkey like T, with a lower status, and some obsessive-compulsive behavioral traits might perceive the same faces as a threatening. Indeed, higher baseline skin conductance measures in monkey T (Mosher et al. 2006, http://ubrp. arizona.edu/conferences/07/abstract.cfm?id=494) suggest a more anxious temperament that might account for his tendency to look less at the *Eye* region. Two opposing factors might be at play when monkeys are confronted with unfamiliar faces: a perceived lower status might explain the reluctance of a monkey to stare at the eyes of another individual; the social attention hypothesis, on the other hand, proposes that individuals with higher status receive more visual attention (Pineda et al. 1994). Insofar as saccades between the two faces in the test pair indicate shifts of visual attention, monkeys H and Q made on average 0.98 and 1.06 midline crossing saccades, respectively, while monkey T, made on average, 4.83 such saccades (Mosher et al. 2006). This increase in the number of saccades might represent furtive looks from one face to the other and suggests a conflict between novelty preference (which holds the eyes of the observer on the novel image) and anxious exploration of a potential social challenge. Regardless of social status; however, the overall results indicate that monkeys are consistent in their ability to use second-order configural processing for faces of conspecifics and in their enhanced ability to recognize faces of their own species. The individual differences, superimposed on these effects, need to be further explored.

References

Bentin S, Sagiv N, Mecklinger A, Friederici A, von Cramon YD (2002) Priming visual face-processing mechanisms: electrophysiological evidence. Psychol Sci 13:190–193

Bruce C (1982) Face recognition by monkeys: absence of an inversion effect. Neuropsychologia 20(5):515–521

Dahl CD, Logothetis NK, Hoffman KL (2007) Individuation and holistic processing of faces in rhesus monkeys. Proc Biol Sci 274(1622):2069–2076

de Haan M, Pascalis O, Johnson MH (2002) Specialization of neural mechanisms underlying face recognition in human infants. J Cogn Neurosci 14(2):199–209

Diamond R, Carey S (1986) Why faces are and are not special: an effect of expertise. J Exp Psychol Gen 115(2):107–117

Dittrich W (1990) Representation of faces in longtailed Macaques (*Macaca fascicularis*). Ethology 85:265–278

Dufour V, Pascalis O, Petit O (2006) Face processing limitation to own species in primates: a comparative study in brown capuchins, Tonkean macaques and humans. Behav Processes 73(1):107–113

Emery NJ (2000) The eyes have it: the neuroethology, function and evolution of social gaze. Neurosci Biobehav Rev 24(6):581–604

Fagan JF (1972) Infants' recognition memory for faces. J Exp Child Psychol 14(3):453–476

Fantz RL (1964) Visual experience in infants: decreased attention to familiar patterns relative to novel ones. Science 146:668–670

Gauthier I, Curran T, Curby KM, Collins D (2003) Perceptual interference supports a non-modular account of face processing. Nat Neurosci 6(4):428–432

Ghazanfar AA, Santos LR (2004) Primate brains in the wild: the sensory bases for social interactions. Nat Rev Neurosci 5:603–616

Ghazanfar AA, Nielsen K, Logothetis NK (2006) Eye movements of monkey observers viewing vocalizing conspecifics. Cognition 101:515–529

Goffaux V, Hault B, Michel C, Vuong QC, Rossion B (2005) The respective role of low and high spatial frequencies in supporting configural and featural processing of faces. Perception 34(1):77–86

Gothard KM, Erickson CA, Amaral DG (2004) How do rhesus monkeys (*Macaca mulatta*) scan faces in a visual paired comparison task? Anim Cogn 7(1):25–36

Guo K, Mahmoodi S, Robertson RG, Young MP (2006) Longer fixation duration while viewing face images. Exp Brain Res 171:91–98

Guo K (2007) Initial fixation placement in face images is driven by top-down guidance. Exp Brain Res 181:673–677

Gunderson V, Swartz KB (1985) Visual recognition in infant pigtailed macaques after 24-hour delay. American Journal of Primatology 8:259–264

Hills PJ, Lewis MB (2006) Reducing the own-race bias in face recognition by shifting attention. Q J Exp Psychol (Colchester) 59(6):996–1002

Kanwisher N, Tong F, Nakayama K (1998) The effect of face inversion on the human fusiform face area. Cognition 68(1):B1–11

Keating C, Keating EG (1993) Monkeys and mug shots: cues used by rhesus monkeys (*Macaca mulatta*) to recognize a human face. J Comp Psychol 107:131–139

Kyes RC, Candland DK (1987) Baboon (*Papio hamadryas*) visual preferences for regions of the face. J Comp Psychol 101(4):345–348

Lacreuse A, Herndon JG (2003) Estradiol selectively affects processing of conspecifics' faces in female rhesus monkeys. Psychoneuroendocrinology 28(7):885–905

Leder H, Bruce V (2000) When inverted faces are recognized: the role of configural information in face recognition. Q J Exp Psychol A 53(2):513–536

Marr D (1980) Visual information processing: the structure and creation of visual representations. Philos Trans R Soc Lond B Biol Sci 290(1038):199–218

Maurer D, Grand RL, Mondloch CJ (2002) The many faces of configural processing. Trends Cogn Sci 6(6):255–260

Mondloch CJ, Maurer D, Ahola S (2006) Becoming a face expert. Psychol Sci 17(11):930–934

Mosher CM, Brooks KN, Spitler KM, Zimmerman PE, Wilder T, Gothard KM (2006) Enhanced skin conductance responses elicited by facial expressions with averted gaze in Rhesus macaques. Program No. 72.11.EE11, 2006 Neuroscience Meeting Planner. Society for Neuroscience, Atlanta, Online

Nachson I, Shechory M (2002) Effect of inversion on the recognition of external and internal facial features. ActaPsychol 109:227–238

Nahm F, Perrett A, Amaral D, Albright T (1997) How do monkeys look at faces? J Cogn Neurosci 9:611–623

Nelson CA (1995) The ontomeny of the human memory: a cognitive neuroscience perspective. Dev Psychol 31:723–738

O'Toole AJ, Vetter T, Blanz V (1999) Three-dimensional shape and two-dimensional surface reflectance contributions to face recognition: an application of three-dimensional morphing. Vis Res 39(18):3145–3155

Overman W, Bachevalier J, Turner M, Peuster A (1992) Object recognition versus object discrimination: comparison between human infants and infant monkeys. Behav Neurosci 106(1):15–29

Overman WH Jr, Doty RW (1982) Hemispheric specialization displayed by man but not macaques for analysis of faces. Neuropsychologia 20(2):113–128

Parr LA, Winslow IT, Hopkins WD (1999) Is the inversion effect in rhesus monkey face-specific? Anim Cogn 2:123–129

Parr LA, Heintz M (2006) The perception of unfamiliar faces and houses by chimpanzees: influence of rotation angle. Perception 35(11):1473–1483

Parr LA, Heintz M, Akamagwuna U (2006) Three studies on configural face processing by chimpanzees. Brain Cogn 62(1):30–42

Parr LA, Winslow JT, Hopkins WD, de Waal FB (2000) Recognizing facial cues: individual discrimination by chimpanzees (*Pan troglodytes*) and rhesus monkeys (*Macaca mulatta*). J Comp Psychol 114(1):47–60

Pascalis O, Bachevalier J (1999) Neonatal aspiration lesions of the hippocampal formation impair visual recognition memory when assessed by paired-comparison task but not by delayed nonmatching-to-sample task. Hippocampus 9(6):609–616

Pascalis O, de Haan M, Nelson CA (2002) Is face processing species-specific during the first year of life? Science 296(5571):1321–1323

Pascalis O, Scott LS, Kelly DJ, Shannon RW, Nicholson E, Coleman M, Nelson CA (2005) Plasticity of face processing in infancy. Proc Natl Acad Sci USA 102(14):5297–5300

Phelps MT, Roberts WA (1994) Memory for pictures of upright and inverted primate faces in humans (*Homo sapiens*), squirrel monkeys (*Saimiri sciureus*), and pigeons (*Columba livia*). J Comp Psychol 108(2):114–125

Pineda JA, Sebestyen G, Nava C (1994) Face recognition as a function of social attention in non-human primates: an ERP study. Brain Res Cogn Brain Res 2(1):1–12

Pourtois G, Schwartz S, Seghier ML, Lazeyras F, Vuilleumier P (2005) Portraits or people? Distinct representations of face identity in the human visual cortex. J Cogn Neurosci 17(7):1043–1057

Rakover SS (2002) Featural vs. configurational information in faces: a conceptual and empirical analysis. Br J Psychol 93:1–30

Rhodes G, Hayward WG, Winkler C (2006) Expert face coding: configural and component coding of own-race and other-race faces. Psychon Bull Rev 13(3):499–505

Richmond J, Colombo M, Hayne H (2007) Interpreting visual preferences in the visual paired-comparison task. J Exp Psychol Learn Mem Cogn 33(5):823–831

Rosenfeld SA, Van Hoesen GW (1979) Face recognition in the rhesus monkey. Neuropsychologia 17(5):503–509

Schyns PG, Bonnar L, Gosselin F (2002) Show me the features! Understanding recognition from the use of visual information. Psychol Sci 13(5):402–409

Sekuler AB, Gaspar CM, Gold JM, Bennett PJ (2004) Inversion leads to quantitative, not qualitative, changes in face processing. Curr Biol 14(5):391–396

Sergeant J (1986) Microgenesis of face perception. In: Ellis H, Jeeves M, Newcombe F, Young A (eds) Aspects of face processing. Kluwer, Dordrecht, pp 17–33

Swartz KB (1983) Species discrimination in infant pigtail macaques with pictorial stimuli. Dev Psychobiol 16(3):219–231

Tarr MJ, Gauthier I (2000) FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. Nat Neurosci 3(8):764–769

Tomonaga M (1994) How laboratory-raised Japanese monkeys (*Macaca fuscata*) perceive rotated photographs of monkeys: evidence of an inversion effect in face perception. Primates 35:155–165

Troje NF, Bulthoff HH (1998) How is bilateral symmetry of human faces used for recognition of novel views? Vis Res 38(1):79–89

Valentine T, Bruce V (1988) Mental rotation of faces. Mem Cognit 16(6):556–566

Vermeire BA, Hamilton CR (1998) Inversion effect for faces in split-brain monkeys. Neuropsychologia 36(10):1003–1014

Vinette C, Gosselin F, Schyns PG (2004) Spatio-temporal dynamics of face recognition in a flash: it's in the eyes. Cogn Sci Multidisciplinary J 28(2):289–301

Wallis G, Rolls ET (1997) Invariant face and object recognition in the visual system. Prog Neurobiol 51(2):167–194

Wright A, Roberts W (1969) Monkey and human perception: inversion effects for human but not for monkey faces and scenes. J Cogn Neurosci 8:278–290

Yin R (1969) Looking at upside-down faces. Exp Psychol 81:141–145

Yovel G, Kanwisher N (2004) Face perception: domain specific, not process specific. Neuron 44(5):889–898