Springer

Dear Author:

Please find attached the final pdf file of your contribution, which can be viewed using the Acrobat Reader, version 3.0 or higher. We would kindly like to draw your attention to the fact that copyright law is also valid for electronic products. This means especially that:

- You may print the file and distribute it amongst your colleagues in the scientific community for scientific and/or personal use.
- You may make your article published by Springer-Verlag available on your personal home page provided the source of the published article is cited and Springer-Verlag and/or other owner is mentioned as copyright holder. You are requested to create a link to the published article in Springer's internet service. The link must be accompanied by the following text: "The original publication is available at springerlink.com". Please use the appropriate DOI for the article. Articles disseminated via SpringerLink are indexed, abstracted and referenced by many abstracting and information services, bibliographic networks, subscription agencies, library networks and consortia.
- Without having asked Springer-Verlag for a separate permission your institute/your company is not allowed to place this file on its homepage.
- You may not alter the pdf file, as changes to the published contribution are prohibited by copyright law.
- Please address any queries to the production editor of the journal in question, giving your name, the journal title, volume and first page number.

Yours sincerely,

Springer-Verlag

Anim Cogn (2004) 7 : 25–36
DOI 10.1007/s10071-003-0179-6

ORIGINAL ARTICLE

**Katalin M. Gothard · Cynthia A. Erickson**
**David G. Amaral**

# How do rhesus monkeys (*Macaca mulatta*) scan faces in a visual paired comparison task?

Received: 24 July 2002 / Revised: 17 March 2003 / Accepted: 19 May 2003 / Published online: 9 July 2003


**Abstract** When novel and familiar faces are viewed simultaneously, humans and monkeys show a preference for looking at the novel face. The facial features attended to in familiar and novel faces, were determined by analyzing the visual exploration patterns, or scanpaths, of four monkeys performing a visual paired comparison task. In this task, the viewer was first familiarized with an image and then it was presented simultaneously with a novel and the familiar image. A looking preference for the novel image indicated that the viewer recognized the familiar image and hence differentiates between the familiar and the novel images. Scanpaths and relative looking preference were compared for four types of images: (1) familiar and novel objects, (2) familiar and novel monkey faces with neutral expressions, (3) familiar and novel inverted monkey faces, and (4) faces from the same monkey with different facial expressions. Looking time was significantly longer for the novel face, whether it was neutral, expressing an emotion, or inverted. Monkeys did not show a preference, or an aversion, for looking at aggressive or affiliative facial expressions. The analysis of scanpaths indicated that the eyes were the most explored facial feature in all faces. When faces expressed emotions such as a fear grimace, then monkeys scanned features of the face, which contributed to the uniqueness of the expression. Inverted facial images were scanned similarly to upright images. Precise measurement of eye movements during the visual paired comparison task, allowed a novel and more quantitative assessment of the perceptual processes involved the spontaneous visual exploration of faces and facial expressions. These studies indicate that non-human primates carry out the visual analysis of complex images such as faces in a characteristic and quantifiable manner.

**Keywords** Face recognition · Facial expression · Inversion effect · Non-human primates · Eye movements

K. M. Gothard · D. G. Amaral
Department of Psychiatry, University of California Davis,
2230 Stokton Blvd., Sacramento, CA 95817, USA

K. M. Gothard · C. A. Erickson · D. G. Amaral
California Regional Primate Research Center,
County Rd. 98, Davis, CA 95616, USA

C. A. Erickson
ARL-Neural System Memory and Aging,
The University of Arizona, 384 Life Sciences North Bldg.,
Tucson, AZ 85724, USA

D. G. Amaral
Center for Neuroscience,
1455 Newton Court, Davis, CA 95616, USA

D. G. Amaral
The M.I.N.D. Institute,
4860 Y Street, Sacramento, CA 95817, USA

*Present address:*
K. M. Gothard
Department of Physiology, College of Medicine,
The University of Arizona, 329 Life Sciences North Bldg.,
Tucson, AZ 85724, USA
Tel.: +1-520-6261448, Fax: +1-520-6261448,
e-mail: kgothard@email.arizona.edu

## Introduction

Recognizing an individual's face or a stereotypical facial expression is effortless, rapid and robust. We tested proficiency of non-human primates in facial recognition in the laboratory using the visual paired comparison task. This task takes advantage of the incidental nature of visual learning and exploits the subject's spontaneous attraction to novelty (Gunderson and Swartz 1985; Gunderson et al. 1988, 1989; Pascalis et al. 1998; Pascalis and Bachevalier 1999; Buffalo et al. 1999; Zola et al. 2000). In this task, a pair of identical images – the familiarization pair – is presented first. After a short delay, a second pair of images – the test pair – is then displayed. One of the images in the test pair had been seen in the familiarization pair and the other image is novel. Preference for the images can be measured as the time spent viewing the novel image compared with the time spent viewing the familiar image. Recognition of the familiar image is indicated by longer looking at the novel image (Fagan 1972). Little work has been done, however, to determine precisely what the subject is attending to during the evaluations of the novel and familiar stimuli. Likewise, little is known about the perceptual processes involved in discriminating between different facial expressions of the same individual. The present study sought to fill these gaps by analyzing the visual search patters (scanpaths) in monkeys performing a visual paired comparison task using faces and facial expressions as visual stimuli.

Monkeys, like humans, rapidly recognize the faces of individuals they encounter (Altman 1967; De Waal 1989; Cheney and Seifarth 1990). Inverting the image of a face has been reported to slow down face recognition in humans and chimpanzees (Diamond and Carey 1986; Bartlett and Searcy 1993; Parr et al. 1998). This effect was interpreted as suggesting the existence of a specialized neural system used for face processing (Diamond and Carey 1986; Tanaka and Farah 1991; Farah et al. 1995; Kanwisher 2000, but for an alternate interpretation see Tarr and Gauthier 2000). Studies investigating how inverted images affect face recognition in rhesus monkeys have shown conflicting results (Rosenfeld and van Hoesen 1979; Bruce 1982; Swartz 1983; Dittrich 1990, Keating and Keating 1993; Tomonaga et al. 1993; Phelps and Roberts 1994; Wright and Roberts 1996; Vermeire and Hamilton 1998; Parr et al. 1999). One objective of the current study was to contribute to the resolution of the controversy over the presence or absence of inversion effect in rhesus monkeys.

Rhesus monkeys communicate their emotions and intentions using a series of characteristic facial expressions (van Hoff 1967; Chevalier-Skolnikoff 1973; Redican 1975). Rapid differentiation and evaluation of the intensity of agonistic and affiliative facial expressions is a survival imperative in many primate societies. Scanpath analyses reveal the order and the extent to which different parts of the face and head are explored during spontaneous exploration of images of facial expressions. Humans explore most extensively the eyes in a face image (Yarbus 1967). We have attempted to determine whether it is the eyes or some other facial feature that commands the attention of a monkey when it is presented with the opportunity to scan familiar and novel images of other monkeys.

This study combines the visual paired comparison task with scanpath analysis to answer four related questions concerning face processing in rhesus monkeys:

1. What is the magnitude of novelty preference for faces compared to non-face stimuli?
2. Which facial features are attended to while comparing novel and familiar faces?
3. Do rhesus monkeys show an inversion effect?
4. Do rhesus monkeys scan facial expressions differently compared to neutral faces of the same individual?

## Methods

### Test subjects

Four male rhesus monkeys (*Macaca mulatta*, 5–14 years old) served as viewer monkeys. All four monkeys were naïve to behavioral testing and were unfamiliar with the monkeys used as face stimuli. The monkeys were born and raised in the outdoor enclosures of the California Regional Primate Research Center. All experimental procedures were performed in compliance with the guidelines of the National Institutes of Health regarding the use of primates in research and were approved by the UC Davis Animal Use and Care Advisory Committee. Each monkey was fitted with a head fixation device attached surgically to the skull under Isoflurane anesthesia. To monitor eye movements, scleral search coils were implanted surgically following standard procedures (Robinson 1963). The subjects were trained under a mild fluid restriction regime (not less than 38 ml/kg average daily water intake).

### Capturing monkey facial expressions on videotape

One of the main components of this project was to develop a library of different facial expressions to use as stimuli for examining monkey social behavior in controlled environments. We attempted to capture on film the behavioral displays or “expressions” that are part of the normal monkey vocabulary of social interaction. By examining conspecific behavioral responses to these expressions, it is possible to infer the meaning of these displays in the context of social interactions. Typical agonistic facial expressions observed in rhesus monkeys include: stare, round-mouth stare, open-mouth stare or open-mouth threat, and bared-teeth stare (Hinde and Rowell 1962; van Hoff 1967). Higher-ranking individuals within a social group are more likely to display these types of facial expressions than lower-ranking individuals. Lower-ranking animals show submissive behaviors or retreat when confronted with agonistic expressions from higher-ranking individuals. An agonistic facial expression may be followed with an aggressive act if the threatened animal does not respond submissively or does not retreat. Facial displays are interpreted in the context of specific social situations, for example, a yawn might become a covert form of threat, especially when a large, dominant male reveals his large canines and strong jaws.

Affiliative or appeasing facial expressions include fear grimace, lipsmack, teeth chatter, and bared-teeth grin (Hinde and Rowell 1962; van Hoff 1967 ). A fear grimace is an appeasing facial expression that indicates subordination and is displayed in response to a threat or the approach of a dominant individual. The most affiliative facial expression is the lipsmack, typically observed during a friendly approach, before grooming, or as a begging gesture toward other monkeys or caretakers. A relaxed face or neutral facial expression is generally seen during resting, grooming or playful activities (Hinde and Rowel 1962; Chevalier-Skolnikoff 1973; Redican 1975).

Fifty-three adult rhesus monkeys (23 females and 30 males) served as models or stimulus monkeys for the test images. On each day of filming a single stimulus monkey was transported into the filming room and placed in a large cage with a transparent front wall. The filming room was 256.5 cm wide and 406.4 cm long. The cage was 81.3 cm wide and 91.4 cm high and deep. It sat 20.3 cm off the floor. Each of the sides had bars, except the front that was a sheet of 0.64 cm Plexiglas. Except for the cage and a dark curtain placed 1.5–2 m from the front of the cage the room was empty from the perspective of the monkey. The curtain (a piece of brown felt draped over a horizontal rod (measured 119.4 cm high and 165.1 cm wide)) prevented the monkey from seeing the video camera and the taping crew. The camera lens protruded through a slit in the curtain. The camera was positioned on a tripod about 91.4 cm off the ground so that it was approximately at the level of the monkey’s head. The taping crew consisted of two to four people, who were initially hidden behind the curtain.

During the first 5–10 min in the cage, the monkey was unaware of the presence of others in the room, and spontaneously explored the unfamiliar environment of the laboratory. From the footage recorded during these initial periods neutral facial expressions were extracted. A crewmember then attempted to induce agonistic and affiliative facial expressions from the monkey while the monkey looked at the camera. To do this, the crewmember protruded his or her head through the slit in the curtain just below or above the camera lens. The monkey usually responded to the sudden appearance of a human face with either an appeasing or a threatening facial expression. To increase the intensity of the monkey's response, the crewmember often mimicked appeasing or threatening monkey facial expressions. If the monkey did not respond to these gestures, a crewmember waved stuffed animals, rubber snakes, etc., in front of the monkey and this often induced fearful facial expressions. Showing the monkeys preferred foods and treats often elicited appeasing gestures, such as lipsmacks. Some monkeys reciprocated human imitations of lipsmacks. It was often necessary to have two crewmembers interact with the monkey. One crewmember would only exhibit those types of behavior observed in high-ranking individuals and a second crewmember would offer treats and show only affiliative behaviors towards the monkey (like those observed in low-ranking individuals). When both moved in front of the curtain and interacted with the monkey, the monkey often displayed threatening expressions towards the threatener and affiliative expressions toward the other crewmember.

## Extracting still images from the videotape

The video footage was recorded with a digital camera (Canon, GL1) on Sony mini premium digital cassettes set to maximum resolution. The camera had a 100× digital zoom lens with optical image stabilization. A digital video editing system using Final Cut Pro software was used to view the footage frame by frame and to extract frames that contained the best examples of monkey facial expressions. The goal was to obtain at least one neutral face, at least two antagonistic, and two affiliative facial expressions from each stimulus monkey (criteria for categorizing facial expressions are shown in Table 1).

Images were selected only if the monkey's face was not obstructed, and the facial expression unambiguously fit into one of the five expression types. Frames containing facial expressions of interest were saved as static image files in tag image file format (TIFF) format, and were further processed in an image-editing program (Adobe Photoshop 6.0). The original static image had dimensions of 720×480 pixels and a resolution of 72 pixels/inch. The image was cropped so that each image contained a monkey face of similar size in its center. The monkey faces took up approximately 75% of the image. Scratches or spots on the Plexiglas wall of the holding cage, as well as small irregularities in color or luminance, were edited out. The final images were saved as 300× 300 pixel Windows bitmap files (BMP). From 60 h of videotape, over 5,000 static images were extracted. In some of these images, the monkey is looking into the camera straight at the viewer. The final images were catalogued according to monkey, facial expression, and gaze direction. On average, 90 images were obtained from each monkey. The entire library contains 1,285 images of threat, 359 images of yawn, 2,452 neutral faces, 183 images of fear grimace and 517 images of lipsmacks. Examples of multiple facial expressions obtained from the same monkey are shown in Fig. 1.

## Training

During training and recording, the monkey was seated in a primate chair in front of a 17-inch computer monitor (Multisync LCD 1525 V) placed 57 cm from the monkey's eyes. At this distance, 1 cm on the monitor corresponds to 1° of visual angle. The chair and the monitor were enclosed in a sound-attenuated booth (Acoustic Systems, Austin, Tex.). Initially, each monkey was trained to orient to, and maintain gaze (fixate) at, small white rectangles that subtended 1° of visual angle in the center of the computer monitor. Monkeys were rewarded for correct responses with a drop of fruit juice paired with a click from the juice delivery system. Errors were followed by 2-s time-out period, in which the screen remained blank. When fixation training was complete, each monkey was then trained to maintain its gaze within the boundaries of a rectangular landscape image that subtended 24×12° of visual angle. Performance criterion was met when monkeys fixated for at least 200 ms on the fixation points and maintained gaze within the boundary of the landscape image for 3 s on 90% of the trials. Each monkey learned the task within 4 weeks of daily training.

## Behavioral task

After initial shaping of looking behavior, the monkeys were trained to look at images in a visual paired comparison task. This task consists of the sequential presentation of two pairs of images with a delay imposed between them. The image pairs were two images located horizontally adjacent to each other. The junction of the image pairs was centered on the monitor. Each image was 12×12° of visual angle so that when they were displayed side-by-side the pair subtended 24×12°. The first pair of images was identical (the familiarization pair). After a randomly selected delay ranging between 1 and 1.5 s, a second pair of images was presented (the test pair). One image in the test pair was the same as in the familiarization pair, whereas the other image was novel. Each trial was performed twice, once with the novel image on the right and once with the novel image on the left. Balancing the position of the each novel image on the left and on the right, eliminated the probability of spurious results due to a possible looking bias. A set of stimuli comprised either 20 object trials or 16 face trials. The sequence of trials within a set was randomized. For each monkey, all five experiments were carried out in one or two data collection sessions.

When the monkeys first saw the face stimuli, especially the aggressive and appeasing facial expressions, they responded to the images with lipsmacks, barks and vigorous ear movements. The head restraint and the task demands limited their motor responses. Behavioral responses to the test stimuli habituated within the first 100 presentation of face images.

## Detailed task protocol

A schematic of the task is shown in Fig. 2. During an experimental session, a white fixation spot (0.25° of visual angle) appeared in

**Table 1** Criteria for classification of facial expressions

| Expression | Eyes | Mouth | Ears | Example |
|---|---|---|---|---|
| Fear grimace | Eye contact, brows raised | Lips retracted, bared teeth, jaws closed, horizontal mid-facial folds | Pulled back | Fig. 8B third image from left |
| Lipsmack | Eye contact, brows raised | Lips puckered or slightly parted, tongue slightly protruded | Pulled back | Fig. 8A, third image from left |
| Neutral | Indifferent | Closed, relaxed | Relaxed | Fig. 8, all the leftmost images |
| Yawn | Indifferent | Widely open, lips retracted, bared teeth | Pulled back | Fig. 8D, fourth image from left |
| Threat | Eye contact | Open, lips cover the upper teeth | Pulled back or forward | Fig. 8C, fourth image from left |

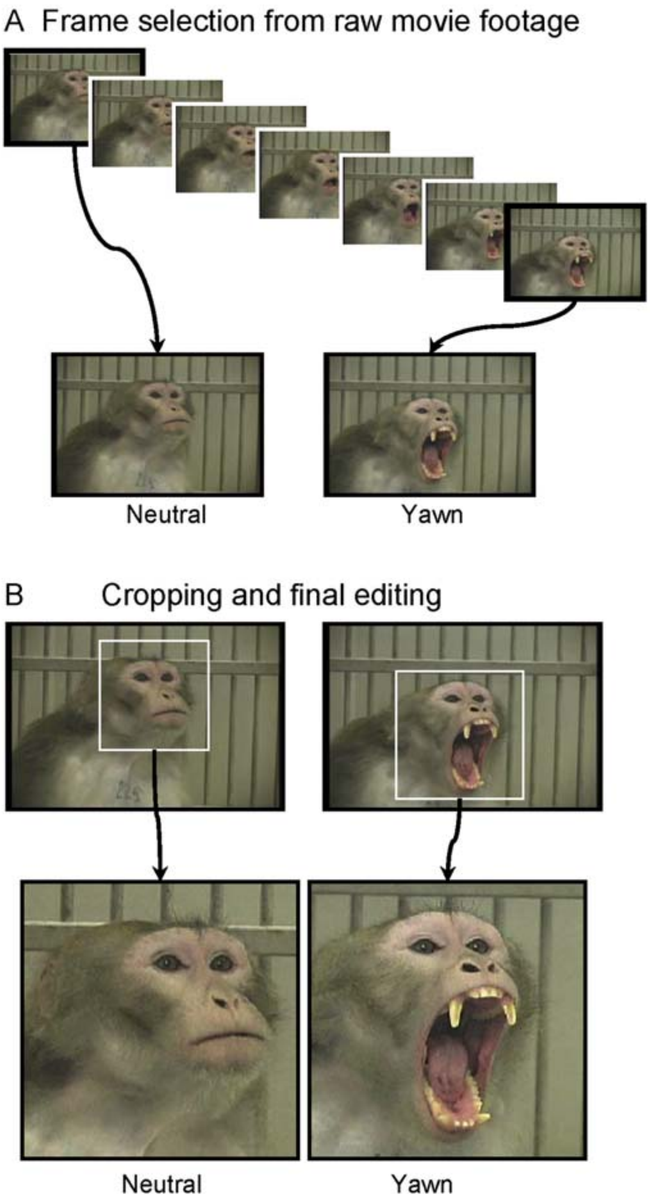


**Fig. 1A, B** This figure shows the process of converting the videotapes of monkey behavior to still images of different facial expressions. **A** The development of a yawn shown frame-by-frame. The monkey starts with a neutral expression and then gradually opens his mouth to yawn. Two frames are selected as representative facial expressions and labeled *neutral* and *yawn*(highlighted in *black)*. **B** Images selected as representative for a particular stereotypical facial display were cropped (*white* rectangles) into 300× 300 pixel images and checked for distracting flaws. The cropped images show the monkey's face centered and occupying approximately 75% of the total image

the center of the screen. The monkey was required to fixate the spot within 1 s of its appearance on the monitor and maintain fixation for at least 250 ms (fixation). If these criteria were not met, a time-out period of 2 s was imposed. If fixation occurred, the familiarization pair was displayed for 1.5 s (familiarization). While the familiarization pair was displayed on the monitor, the monkey was free to explore the two images without shifting its gaze outside their boundary. If the monkey's gaze moved outside the boundary of the image pair, the screen went blank for 2 s, signaling failure on the trial. If, however, the animal maintained its gaze within the image boundary for 1.5 s, the image disappeared and a drop of fruit juice was delivered into its mouth as a reinforcer.

After a delay of 1 s, plus a random delay of up to 500 ms (delay), the fixation spot reappeared (re-fixation). If the monkey did not 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 1.5 s (test). Again, the monkey was required to maintain its gaze within the boundary of the image pair for the duration of the display. At the completion of the trial, the monkey received two drops of juice reward. The juice reward was paired with an auditory stimulus providing the animals with an indication of the correct response. After an inter-trial interval of 2 s, the next trial began with the display of the fixation spot. Error trials were repeated later in the experimental session.

## Measurement of eye movements

Eye movements were tracked with a resolution of 0.25° of visual angle and sampled and digitized at 500 Hz using an eye tracker (DNI, Newark, Del.) and the CORTEX data acquisition program (kindly provided by Dr. Robert Desimone, The Laboratory of Neuropsychology, NIMH). Scanpaths are defined as digitized eye movements calibrated in degrees of visual angle and superimposed on stimulus images. For each experiment, novelty preference was quantified as the amount of time (number of points or samples) spent looking at either the novel or the familiar image. A point was a sample of eye position in Cartesian coordinates, wherever the eye was pointed within the boundaries of the image. Novelty preference was expressed as percent time spent looking at either the novel or familiar images relative to the total looking time. In order to quantify the viewing time of facial features, an experimenter, blind to the monkey's scanpaths, drew rectangles around the facial features with a mouse on a computer screen (the rectangles are shown in Figs. 4, 6, and 7). These rectangles represent regions of interest and included a facial feature (e.g., the ear) with the margins slightly exceeding the feature to ensure that a fixation on the margin of that feature (e.g., looking at the tip of ear) will be counted as looking at that feature. Because different monkeys have differently sized facial features, these regions of interest also varied in size with the individual. Most importantly, all the regions of interest were drawn by the same experimenter before the scanpath was overlaid on the image. When the scanpath was overlaid on the stimulus face, the looking time was quantified using Matlab (Math Works, Inc., Mass.) routines. These routines added all the points that fell within the boundaries of the regions of interest. Although we do not process visual information during saccades, adding the samples recorded during saccades to the samples recorded during fixations does not introduce significant fluctuations in the data because fixations last two orders of magnitude longer than saccades. Looking time within and outside the regions of interest added up to 100%. Statistical analysis of looking time was performed using one-way and two-way repeated measures ANOVA.

## Stimulus sets

For each of the four experimental questions, different stimulus sets were presented to the animal. The stimulus set for experiment 1 consisted of images of 20 junk objects (i.e., 10 familiar and 10 novel objects). Ten objects were assigned for the familiarization pairs (two identical images) and the other 10 objects for the novel half of the test pairs. Each of the 10 novel objects was presented twice, with 10 novel object on the right and on the left of the test pair, thus the object session consisted of scanpath on 40 image pairs. The stimulus set for experiment 2, comprised 16 images of neutral monkey faces, of which 8 images were assigned to the familiarization pairs and another 8 images to be the novel face in the test pair. Again, each image pair was presented twice, with the novel face on the left and on the right, thus the monkey performed a total of

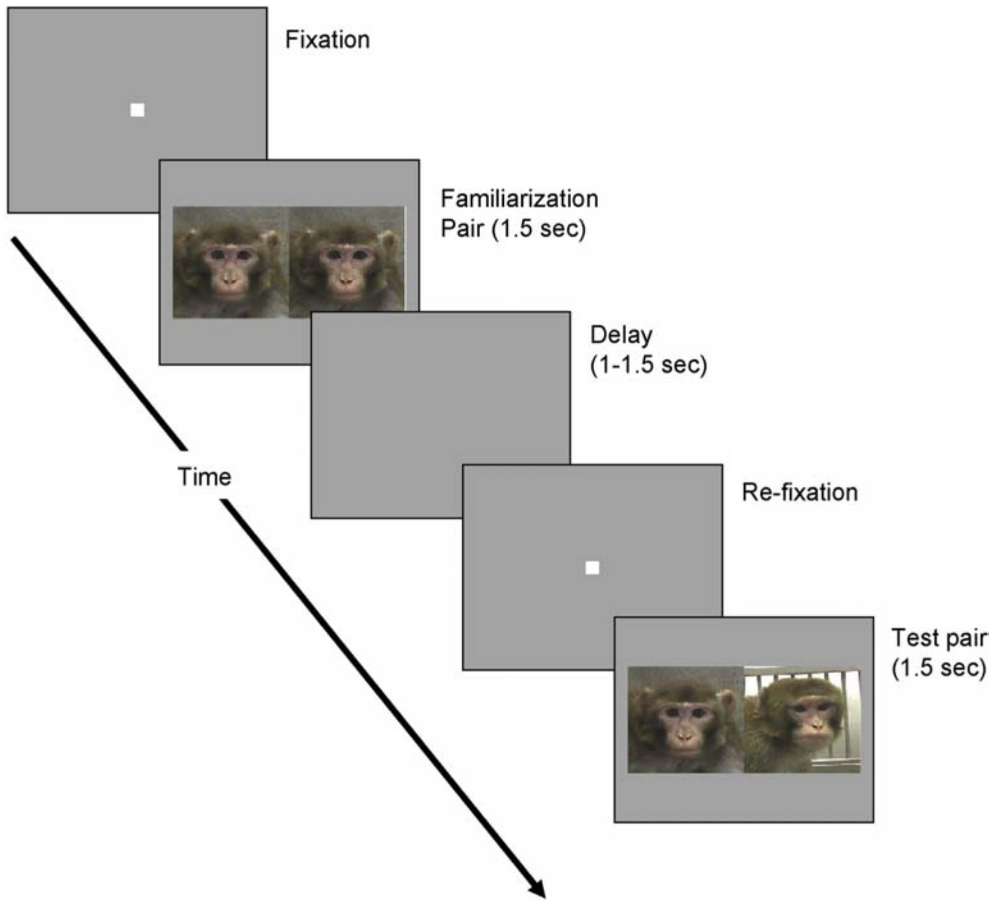


**Fig. 2** Schematic of the visual paired comparison task. The five partially overlapping panels represent the appearance of the display monitor during each phase of the task. The *arrow* indicates time. The first fixation spot (upper left panel, *fixation*) was displayed for 1.5 s or until the monkey looked at the spot for 250 ms. The second panel shows an example of a familiarization pair (*familiarization pair*), consisting of two juxtaposed identical images. The familiarization pair was displayed for 1.5 s. The third panel shows a blank screen (*delay*) of 1–1.5 s. The fourth panel shows the presentation of the second fixation spot (*re-fixation*). The lower right panel shows the corresponding test pair image (*test pair*). The image on the left has been seen in the familiarization pair, the image on the right is novel. Novel and familiar images were randomly counterbalanced for spatial location

36 trials (counterbalanced design). The same 16 images were used in experiment 3 but the images were inverted. The stimulus set for experiment 4 consisted of upright neutral images and images of four different facial expressions taken from the same eight monkeys (two monkeys with each expression). In the images used for the familiarization pairs all monkeys displayed a neutral expression. For the eight test pairs, the monkeys displayed lipsmack, fear grimace, open-mouth threat, and yawn expressions.

## Results

Data were analyzed for 1.5 s of image viewing time. In this interval, the monkeys made 5.6±0.75 (mean±SD) saccades. All data presented here were recorded from correct and completed trials.

### Experiment 1: effects of experience on the looking patterns elicited by objects

The first experiment sought to establish a baseline value of novelty preference by recording the monkeys' looking patterns when shown pairs of object images. We hypothesized that all monkeys would look longer at the novel objects. As expected, each one of the four monkeys showed a clear novelty preference because they explored the novel side of the test pairs longer than the familiar side. On average, the monkeys spent 76% of the time looking at the novel object and 24% viewing the familiar object (Fig. 3). The amount of time looking at the novel object was significantly different from chance [$t(3)=4.8$, $P<0.01$].

### Experiment 2: looking patterns elicited by novel and familiar neutral monkey faces

The goal of the second experiment was to test the hypothesis that monkeys spend more time looking at novel faces

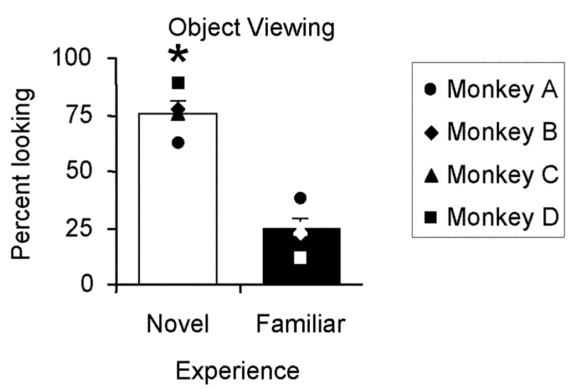


**Fig. 3** Preference for novel objects. Percent looking time (mean± SEM) on *familiar* and *novel* objects. All four monkeys explored the novel object for a longer period of time (the *symbols* indicate time spent viewing by individual monkeys, *asterisk* represents $P<0.01$)

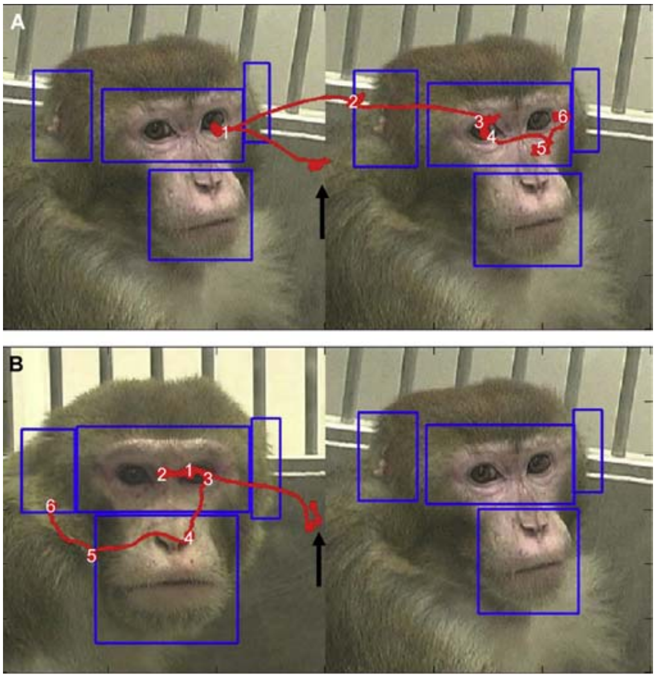


**Fig. 4A, B** Examples of scanpaths recorded from monkeys performing the visual paired comparison task on conspecific faces with neutral facial expressions. The *red lines* correspond to the trajectory of the eyes over the image. The *blue rectangles* represent boundaries centered on the ears, the mouth and the eyes and used to quantify looking time on these features. The scanpath (eye trace) always starts in the center of an image pair, indicated here by the *arrow*. The fixations are numbered in order – see text for full discussion. The majority of fixations are located within the rectangles enclosing the eyes. **A** During the familiarization pair, the monkey looks slightly longer at the face on the right. **B** On the test pair, all the fixations occur on the novel face presented on the left

than at familiar faces. We also determined which facial features are attended to when monkeys scan familiar compared to novel faces. For example, the top panel of Fig. 4A shows the looking pattern (in red) on the two faces of the familiarization pair. The blue rectangles are the regions of interest used to quantify the viewing time of a particular facial feature (eyes, mouth and ears). The viewer monkey is looking at the center of the screen when the image pair appears on the screen (this initial fixation is indicated by the black arrow). Then, it makes an eye movement, to the eyes of the monkey shown on the left panel (1). His gaze rests there for approximately 200 ms, and then he makes a saccade to near the ear of the monkey in the right panel (2), and for the rest of the trial looks almost exclusively at or near the eyes of that face (3–6). The eye movements, or scanpath, overlaid on the test pair of faces is shown in the lower panel (Fig. 4B). The first three saccades (1–3) were directed towards the eyes of the novel monkey face. The next saccade was directed towards the nose (4), then the cheek or side of the face, which falls in the category "other" (5), and finally the ear (6). Note that in this example the viewer monkey looked exclusively at the novel face.

These observations were quantified by calculating the number of time points in which the monkey spent viewing

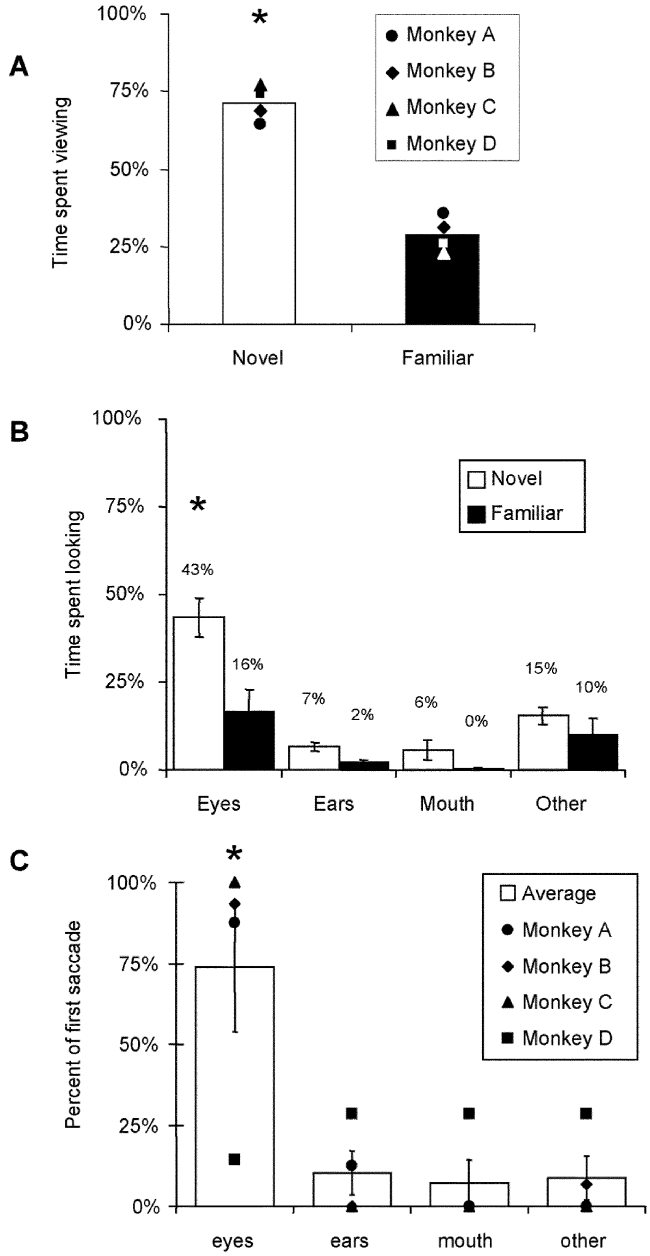


**Fig. 5A–C** Novelty and feature preference for neutral monkey faces. **A** Mean looking time on the novel and familiar face in the test pair. The *symbols* indicate viewing preference from each monkey. Monkeys explored the novel face for significantly longer periods of time. **B** The distribution of looking time by facial features was expressed as percent time when the gaze was inside the boundaries of the regions of interest drawn around the eyes, the ears, the mouth and the regions of the face not enclosed by rectangles (*other*). Monkeys explored the eyes significantly longer than any other regions of the face. The eyes in the novel face were explored longer than the eyes of the familiar face. On familiar faces, no difference was found between time spent exploring the eyes and the regions of the image not enclosed by rectangles. **C** The target of first saccade was the eye area of the stimulus monkey for three of the four monkeys

either the novel or familiar faces (familiarization effect) and the number of time points spent viewing different facial features defined by the boundaries of the regions of interest (shown as blue rectangles in Fig. 4). The relationship between experience level (two levels: novel and familiar) and facial features (four levels: eyes, ears, mouth, and other) was examined with a 2-way repeated measures ANOVA. There was a significant interaction between experience level and facial features [$F(3,24)=3.6$, $P<0.05$]. There were two significant main effects: (1) monkeys looked significantly longer at the novel face [$F(1,24)=13.8$, $P<0.005$, Fig. 5A] and (2) they differentially explored the facial features [$F(3,24)=19.7$, $P<0.001$, Fig. 5B]. Post hoc analyses revealed that this interaction is explained by preferential viewing of eyes in the novel face (Fig. 5B).

We quantified the targets of the first saccades on all the novel neutral faces (Fig. 5C) and it appears that in the majority of cases the target of the first saccade on the novel face was to the eye area of the face. There was a significant main effect of facial feature [$F(1,12)=7.8$, $P<0.005$] despite the observation that only three of the four viewer monkeys directed their first saccade towards the eye of the stimulus faces.

## Experiment 3: face recognition and looking patterns elicited by novel and familiar inverted faces

This experiment sought to determine whether monkeys recognize inverted faces and whether inverted faces elicit looking patterns similar to those exhibited for upright faces.

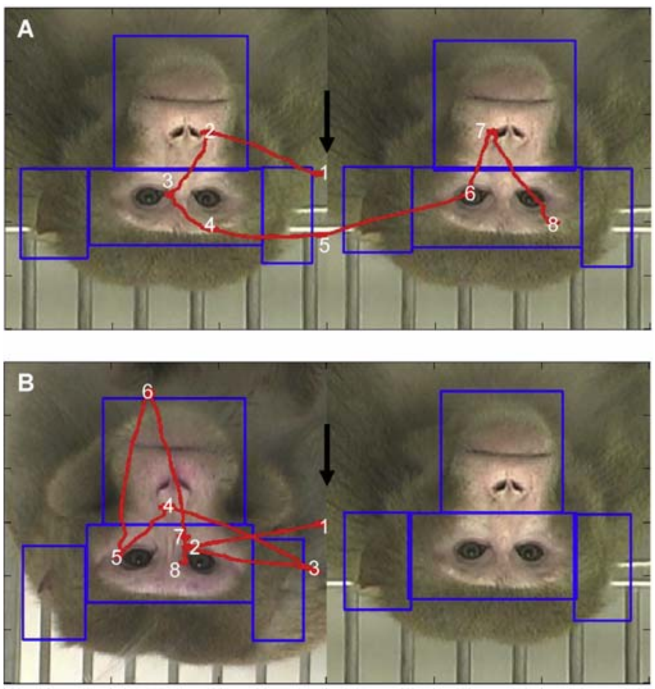


**Fig. 6A, B** Example of scanpaths recorded from Monkey A performing the visual paired comparison task with inverted neutral faces. **A** During presentation of the familiarization pair, the monkey scans both faces and looks at the eyes and nose of the inverted monkey face. **B** When the test pair was presented, the viewer monkey exclusively scanned the novel face and fixated primarily on the eyes

We hypothesized that rhesus monkeys show no inversion effect and therefore, they look longer at the novel faces even if presented upside-down (Fig. 6).

In the example shown in Fig. 6, the scanpath (in red) on the familiarization pair (top panel), indicates the initial fixation (1) the viewer monkey looked at the area of the mouth, specifically the nose, of the left inverted face (2). Following along this scanpath, the third and fourth (3 and 4) fixations targeted the eyes. The fifth fixation (5) occurs

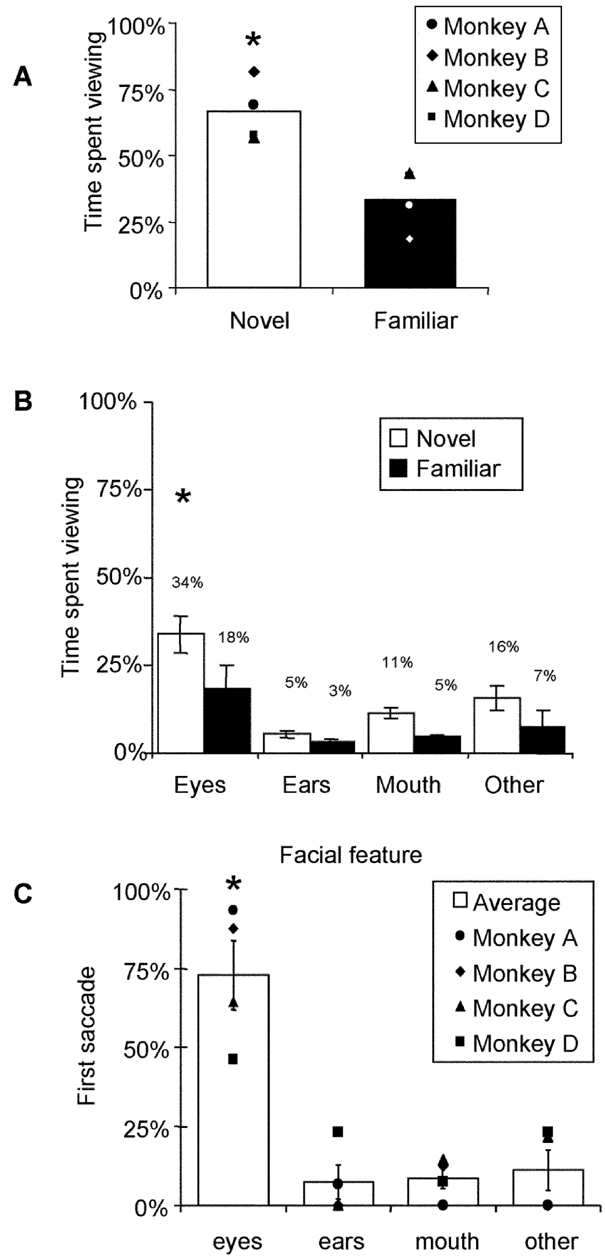


**Fig. 7A–C** Novelty and feature preference for inverted faces. **A** The mean looking time on familiar and novel inverted faces, expressed as percent of total looking time on the test pair. **B** The mean (±SEM) looking time on ears, eyes, mouth and other regions of the face expressed as percent of the total looking time on the familiar and novel face. **C** The target of the first saccade during test pair presentation

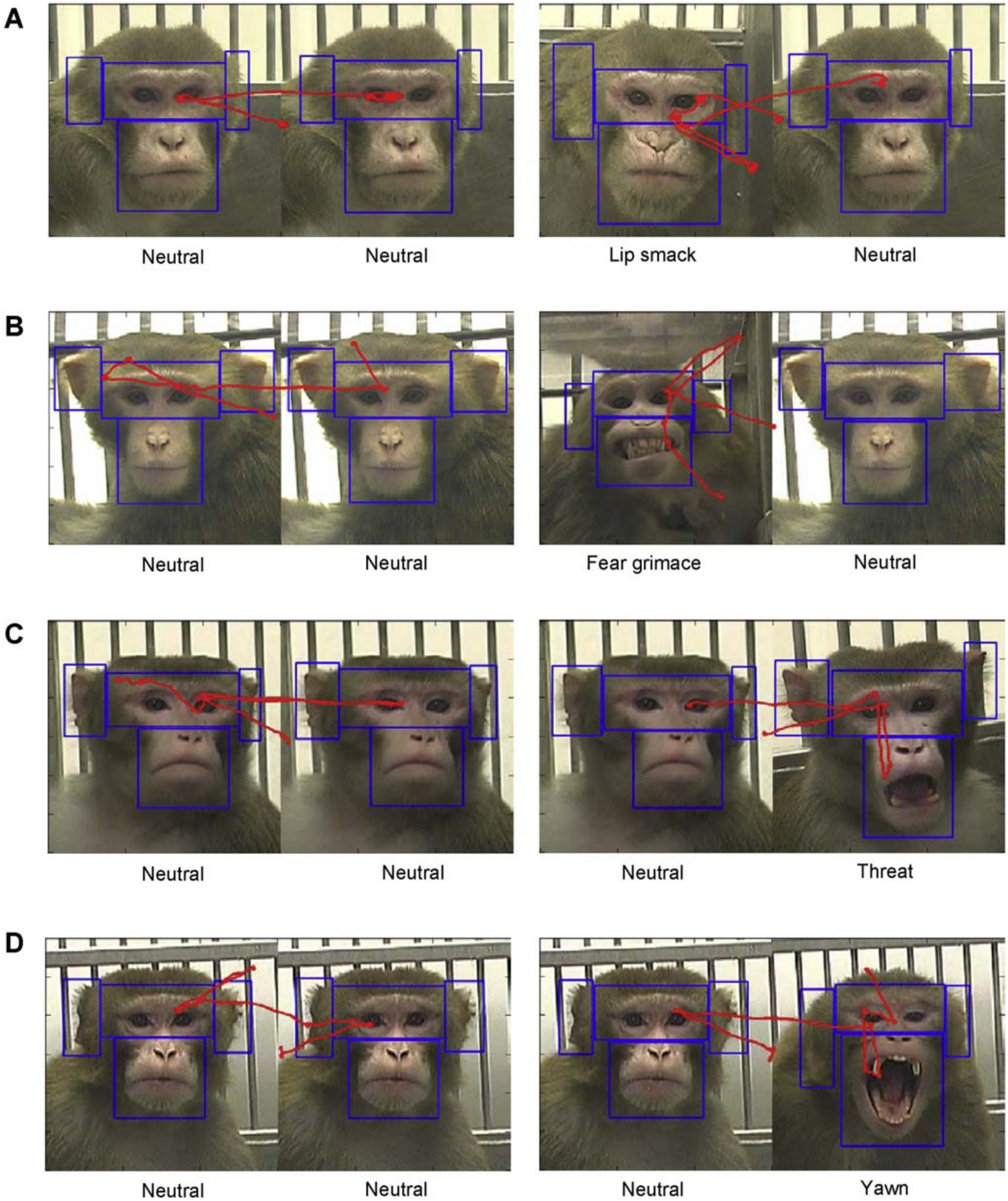


**Fig. 8A–D** Examples of scanpaths recorded from monkeys performing the visual paired comparison task with novel facial expressions. Each row represents a trial. In the familiarization pair the stimulus monkey appears with a neutral face. In the test pair, the same stimulus monkey is shown with a neutral face and an aggressive or affiliative facial expression. The four facial expressions are *fear grimace* (**A**), *lipsmack* (**B**), open-mouth *threat* (**C**) and *yawn* (**D**). Note that the viewer monkey fixated at the canines where visible (**B** and **D**) on the boundary between the two images. The next saccade (6) targets the eyes followed by a saccade to the nose (7). The last fixation during the 1,500 ms scanning time targeted the eye (8). Note the relative symmetry of the scanpath in the two identical faces, and that on both sides the eyes were the most extensively explored feature. On the test pair (shown in Fig. 5B) the first saccade targets the eyes of the novel face (2). After a short fixation on the boundary between the two images (3) the viewer monkeys fixates at the nose (4). There were three more fixations on the eye area (5, 7, and 8) interrupted by a saccade targeting the chin area, which in this particular case was classified as other because the fixation occurred outside the boundary of the rectangle enclosing the mouth. In this test pair, as in many other test pairs, the familiar images were not explored.

The relationship between experience level (two levels: familiarity and novelty) and facial feature (four levels: eyes, ears, mouth, and other) was examined with 2-way repeated measures ANOVA. While there was no interaction between these two factors, there were two significant main effects. All four monkeys spent more time scanning the novel inverted faces relative to the familiar inverted faces [$F(1, 24)=11.3$, $P<0.003$] (Fig. 7A). As in the case of upright faces, there was a significant difference in viewing time of different facial features [$F(3,12)=20.41$, $P<0.0001$]. Pairwise comparisons revealed that when scanning the inverted monkey faces, the monkeys looked longer at the eyes than the mouth ($P<0.05$) and they looked less at the mouth than either the eyes or the ears (Fig. 7B).

As in the case of upright faces, on inverted faces the first saccade was also directed to the eyes of the face stimuli, preferentially to the eyes of the novel monkey face [$F(1,12)=7.8$, $P<0.005$; Fig. 7C]. We also compared the amount of time looking at the eyes in the neutral condition (from experiment 2) to the amount of time looking at the eyes in the inverted condition. There was no significant difference [$t(3)=1.36$, $P>0.05$].

### Experiment 4: face recognition on image pairs where the novelty is a new facial expression

The fourth experiment was devised to test the hypothesis that monkeys show a significant preference for novel facial expressions of the same individual (Fig. 8). To test this hypothesis we compared looking preference for two pairs of images depicting the same individual. In the first pair of images (familiarization pair) a stimulus monkey with a neutral expression was presented on both sides. The second pair (test pair) consisted of the same neutral face paired with a stereotypical facial expression (fear grimace, lipsmack, open mouth threat, and yawn) produced by the same stimulus monkey. The typical scanpaths (shown in Fig. 8) recorded in this experiment showed that the viewer monkeys explore almost exclusively the eyes of the neutral faces, but attend to other facial features in all four facial expressions. For example, in facial expressions where the mouth is open, exposing the canines (Fig. 8B–D) at least one saccade targets the mouth or specifically the canines (Fig. 8D).

The relationship between facial expression and facial feature was examined with 2-way repeated measures ANOVA. There was a significant interaction between these two factors [$F(9,48) = 20$, $P<0.001$]. The interaction is explained by a preferential viewing of certain facial expressions. There was a significant main effect of experience [$F(3,48)= 3.05$, $P<0.01$], indicating that the monkeys looked longer at the novel facial expression than at the familiar neutral face. There was no main effect of facial expression [$F(3,48)=0.72$, $P>0.05$]. As in the case of neutral faces, the viewer monkeys looked first at certain facial features [$F(3,24)=10.83$, $P<0.01$]. As shown in Fig. 9, when monkeys explored the two aggressive facial expressions (threat or yawn) where the mouth figures prominently in the expression, time was distributed about equally between the eyes and the mouth. Fear grimaces induced the longest looking to the mouth, while lipsmacks induced the longest looking time to the eyes.

## Discussion

Three main findings resulted from these experiments. First, monkeys explored images of novel faces significantly longer regardless of their orientation (upright or inverted) and facial expression. The magnitude of novelty preference for faces and objects was similar and on average was about three times longer looking at novel than at familiar

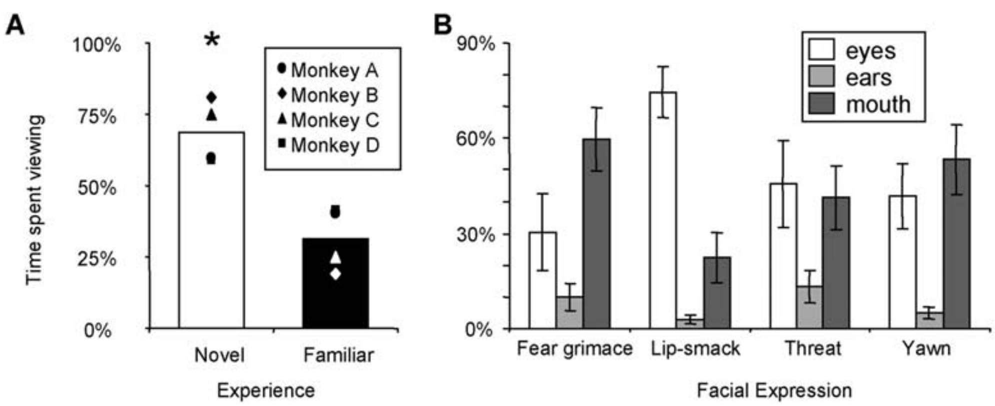


**Fig. 9** **A** The percent of time looking at novel half of the stimulus pair was higher than the looking time at the familiar half. **B** Mean looking time on *ears*, *eyes*, *mouth* and other areas calculated for four facial expressions: *fear grimace*, *lip smack*, *yawn* and *threat*. Looking time on each feature is expressed as percent of total looking time

images. Second, the eyes were the most extensively explored facial feature in faces with neutral expression. Third, the time spent looking at the eyes, mouth, and ears varied across facial expressions.

The values of novelty preference reported here were comparable or larger than reported by earlier studies (Gunderson and Swartz 1985; Gunderson et al. 1988, 1989; Bachevalier et al. 1993; Pascalis and Bachevalier 1999; Buffalo et al. 1999; Zola et al. 2000). This might be the case, because in our version of the task, the monkeys were required to attend to the stimulus images and maintain their gaze within the boundary of the image in order to obtain reward. In all previous studies, novelty preference was based on spontaneous visual scanning, where the subjects were allowed to look away from the stimulus images. The variability in our data was further reduced by high spatial resolution and high-frequency sampling of eye position, compared to scoring videotapes of the viewer monkey's face and eyes. The larger novelty preference reported here may also be accounted for by two additional factors related to task parameters: shorter presentation time and shorter delays between the presentation of the familiarization and the test pairs. Bachevalier et al. (1993) suggested that shorter presentation times might capture more accurately the initial peak of novelty preference. Indeed, in our pilot experiments, when monkeys were allowed to view the image pairs for 3 s or longer novelty preference was more pronounced in the first 2 s of free viewing (data not shown).

One controversial issue of face recognition is the so-called inversion effect. In short, humans and chimpanzees discriminate upright faces more rapidly than inverted faces (Yin 1969; Diamond and Carey 1986; Farah 1995; Farah et al. 1998; Parr et al. 1998). Some studies have found inversion effects in macaque monkeys (Overman and Doty 1982; Swartz 1983; Keating and Keating 1993; Vermeire and Hamilton 1998) while others have not (Rosenfeld and van Hoesen 1979; Bruce 1982; Dittrich 1990; Tomonaga et al. 1993; Parr et al. 1999). The inversion effect has led to the proposition that faces belong to a special category of stimuli that are processed in dedicated neural systems (Diamond and Carey 1986; Tanaka and Farah 1991; Farah et al. 1995; Kanwisher 2000). It is suggested that inverted faces are handled by other neural systems involved in object recognition. The delay in facial discrimination, therefore, may come from the competition of brain systems employed in perceiving inverted faces (Is it a face or an object?).

The monkeys tested in this study did not show impairment in discriminating between inverted familiar and novel faces, suggesting that rhesus monkeys process objects and faces similarly. Given that monkeys appear to discriminate between familiar and novel inverted faces, it was of interest to determine whether they use the same scanpath for viewing inverted versus upright faces. The observations that the distribution of looking time on facial features was similar for upright and inverted faces and that the targets of the first saccade on both upright and inverted faces were the eyes, further strengthen the hypothesis that rhesus monkeys process upright and inverted faces similarly. Task parameters and the choice of face and non-face images appear to influence the outcome of tests for the inversion effect. Parr and colleagues found that rhesus monkeys performed a sequential match-to-sample task significantly better on upright than on inverted images of automobiles, rhesus monkey and capuchin faces, but not human faces or abstract images (Parr et al. 1999). They also found that only humans and chimpanzees, show an inversion effect for stimuli that require expertise for discrimination. The absence of an inversion effect in our study can be accounted for by the different perceptual and cognitive demands of the visual paired comparison task compared to the matching-to-sample task used by Parr and colleagues. On the other hand, the absence of the inversion effect might reflect the monkeys' familiarity with viewing inverted faces. Although rhesus monkeys are not arboreal, they often adopt positions in which the face is held upside-down. For example, lipsmacking with the head upside-down between the hind legs is a common display of submission in rhesus monkeys. These data suggest that monkeys process faces and objects using the same neural pathways. If this is the case, face-selective neurons should be intermixed with neurons selective for objects in the higher visual areas of the rhesus temporal lobe.

The second main finding of this study was that the eyes were the most extensively explored feature in all faces and the eyes were also the targets of the first saccade on the novel face. This suggests that monkeys, like humans, rely heavily on the appearance of the eyes to match the face being viewed with stored memories of previously seen faces. The analysis of scanpaths in humans and in some primate species indicates that all primates tend to spend more time looking at the eyes than at any other facial feature, and rely heavily on the eyes to discriminate between the faces of individuals (Luria et al. 1964; Macworth and Morandi 1967; Yarbus 1967; Walker-Smith et al. 1977; Keating and Keating 1982; Kyes and Candland 1987; Rizzo et al. 1987; Nahm et al. 1998). In both upright and inverted neutral faces, the second most explored facial features fell into the other category. Targets in the other category included the forehead, neck, upper torso, etc., but also areas of the image outside the contour of the monkey, for example, the boundary between images. Recall that subjects were required to fixate on a small icon in the center of the monitor before the stimuli were displayed; hence the first 200–300 ms corresponding to the reaction time before the first saccade was initiated always fell in the other category. Furthermore, the outline of the face and the high-contrast boundary between the two juxtaposed images was often explored, which further increased the amount of time the subjects spent looking at areas other than the rectangles delineating the eyes, ears, and mouth (note saccades targeting the boundary between images in Fig. 6). The third most explored facial feature was the mouth. The same relative importance of facial features was observed by Dittrich (1990) using behavioral measures in freely moving monkeys. Dittrich concluded that for long-tail macaques the main source of information was the outline of the face, followed by the eyes and the mouth.

Our novel approach of measuring the eye movements used to scan faces allowed us to determine not only the facial feature preference across the board but also to explore the possibility that the same facial feature is explored differently when it is contributing to a stereotypical expression. While looking at faces with neutral expressions, facial feature preference was biased in favor of the eyes whereas in emotional facial expressions the distribution of looking time to the eyes, mouth and ears varied with each expression. Our results indicate that images of facial expressions elicit exploration patterns that involve not only the eyes but also those facial features that are the most characteristic for a particular expression. For example, monkeys explored the eyes and the mouth to the same extent on threatening facial expressions (threat and yawn). On yawning faces the saccade to the mouth area was directed precisely to the canines at about 50% of scanpaths (an example is shown in Fig. 8D), while on faces with an open-mouth threat expression, the saccades were directed to the mouth and to the ears. The looking patterns on the two appeasing facial expressions (fear grimace and lipsmack) were different. Monkeys explored predominantly the mouth area on fear grimaces (Fig. 8B). In contrast, images of lipsmacking faces were viewed more like faces with neutral expression, where the eyes were the most extensively explored feature. The conclusion that different facial displays elicit different scanning patterns is in accordance with similar observations on social recognition in monkeys. For example, Dittrich (1994) reported that monkeys use different features of the same stimulus (different body parts or facial features) to differentiate between images of monkeys of different species.

Although we used scanpaths to determine if a face was preferred and if facial perception or recognition is related to a particular pattern of visual exploration, eye movements are not necessary for face recognition. Monkeys trained to fixate on a small icon during the presentation of images have been shown to perform visual recognition tasks accurately (Chelazzi et al. 1993). Eye movements involved in scanning an image likely indicates spontaneous shifts of attention toward the most interesting or informative details (Yarbus 1967, Walker-Smith et al. 1977, Daffner et al. 1999). If this is the case, our results indicate that, the eyes are certainly the most interesting and informative facial feature to monkeys when they attempt to identify and discriminate between conspecifics with neutral expressions. Differential exploration of features in emotional facial expressions may be a useful indicator of the relative importance of various facial features in displays used by monkeys for social communication.

**Acknowledgements** We thank Erin Mitchell for help with animal training, data collection, and image editing. Melody Haworth, GregVicino, Ji-Young Woo and Henry Alitto were members of the videotaping crew. Melody Haworth also helped with image editing. Daniel Bucher wrote programs for data analysis. This study was supported by: K01MH01902, NCRR base grant for the CRPRC, FARDF (Department of Psychiatry, UC Davis, CA), RR00169 MH/MHD5750201A1 and AG03376.

## References

Altman SA (ed) (1967) Social communication among primates. University of Chicago Press, Chicago

Bachevalier J, Brickson M, Hagger C (1993) Limbic-dependent recognition memory in monkeys develops early in infancy. Neuroreport 4:77–80

Bartlett JC, Searcy J (1993) Inversion and configuration of faces. Cogn Psychol 25:281–316

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

Buffalo EA, Ramus SJ, Clark RE, Teng E, Squire LR, Zola SM (1999) Dissociation between the effects of damage to perirhinal cortex and area TE. Learn Mem 6:572–99

Chalazzi L, Miller EK, Duncan J, Desimone R (1993) A neural basis for visual search in inferior temporal cortex. Nature 363: 345–347

Cheney DL, Seyfarth RM (1990) How monkeys see the world: Inside the mind of another species. University of Chicago Press, Chicago

Chevalier-Skolnikoff S (1973) Facial expressions of emotion in non-human primates. In: Ekman P (ed) Darwin and facial expressions. Academic Press, New York, pp 11–90

Daffner KR, Mesulam MM, Cohen LG, Scinto LF (1999) Mechanisms underlying diminished novelty-seeking behavior in patients with probable Alzheimer's disease. Neuropsychiatry Neuropsychol Behav Neurol 12:58–66

De Waal FB (1989) Peacemaking among the primates. Harvard University Press, Cambridge Mass.

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

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

Dittrich WH (1994) How monkeys see others: discrimination and recognition of monkeys' shape. Behav Processes 33:139–154

Fagan JF (1972) Infants' recognition memory for faces J Exp Child Psychol 143:453–76

Farah MJ, Tanaka JW, Drain HM (1995) What causes the face inversion effect? J Exp Psychol Hum Percept Perform 21:628–634

Farah MJ, Wilson KD, Drain M, Tanaka JN (1998) What is "special" about face perception? Psychol Rev 105:482–498

Gunderson VM, Swartz KB (1985) Visual recognition in infant pigtailed macaques after 24-hour delay. Am J Primatol 8:259–264

Gunderson VM, Grant-Webster KS, Burbacher TM, Mottet NK (1988) Visual recognition memory deficits in methylmercury-exposed *Macaca fascicularis* infants. Neurotoxicol Teratol 10: 373–379

Gunderson VM, Grant-Webster KS, Sackett GP (1989) Deficits in visual recognition in low birth weight infant pigtailed monkeys (*Macaca nemestrina*). Child Dev 60:119–127

Hinde RA, Rowel TE (1962) Communications by postures and facial expressions in the rhesus monkey (*Macaca mulatta*). Proc Zool Soc (Lond) 138:1–21

Kanwisher N (2000) Domain specificity in face perception. Nat Neurosci 38:759–63

Keating CF, Keating EG (1982) Visual scan patterns of rhesus monkeys viewing faces. Perception 11:211–219

Keating CF, 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 regins of the face. J Comp Psychol 101:345–348

Luria AR, Pravdina-Vinarskaya EN, Yarbus AL (1964) Eye movement mechanisms in normal and pathological vision. Sov Psychol Psychiatry 2:28–39

Macworth NH, Morandi AJ (1967) The gaze selects informative details within pictures. Percept Psychophys 2:547–551

Nahm FKD, Perrett A, Amaral DG, Albright TD (1998) How do monkeys look at faces? J Cogn Neurosci 9:611–623

Overman WH, Doty RW (1982) Hemispheric specialization displayed by man but not macaques for analysis of faces. Neuropsychologia 20:113–28

Parr LA, Dove T, Hopkins WD (1998) Why faces may be special: evidence of the inversion effect in chimpanzees. J Cogn Neurosci 10:615–622

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

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

Pascalis O, de Haan M, Nelson CA, de Schonen S (1998) Long-term recognition memory for faces assessed by visual paired comparison in 3- and 6-month-old infants. J Exp Psychol Learn Mem Cogn 24:249–60

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

Redican WK (1975) Facial expressions in non-human primates. In: Rosenblum LA (ed) Primate behavior: developments in field and laboratory research vol 4, Academic Press, New York, pp 103–194

Rizzo M, Hurtig R, Damasio AR (1987) The role of scanpaths in facial recognition and learning. Ann Neurol 22:41–45

Robinson DA (1963) A method of measuring eye movements using a scleral search coil in a magnetic field. IEEE Trans Biomed Eng 10:137–145

Rosenfeld SA, Van Hoesen GW (1979) Face recognition in the rhesus monkey. Neuropsychologia 175:503–509

Swartz KB (1983) Species discrimination in infant pigtail macaques with pictorial stimuli. Dev Psychobiol 16:219–31

Tanaka JW, Farah MJ (1991) Second-order relational properties and the inversion effect: testing a theory of face perception. Percept Psychophys 50:367–72

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

Tomonaga M, Itakura S, Matsuzawa T (1993) Superiority of conspecific faces and reduced inversion effect in face perception by a chimpanzee. Folia Primatol Int J Primatol (Basel) 61: 110–4

van Hoff J (1967) The facial displays of the catarrhine monkeys and apes. In: Morris D (ed) Primate ethology, 1st edn. Aldine, Chicago, pp 7–68

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

Walker-Smith GJ, Gale AG, Findlay JM (1977) Eye movement strategies involved in face perception. Perception 63:313–26

Wright AA, Roberts WA (1996) Monkey and human face perception: inversion effects for human but not for monkey faces or scenes. J Cogn Neurosci 8:278–290

Yarbus AL (1967) Eye movements and vision. Plenum, New York

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

Zola SM, Squire LR, Teng E, Stefanacci L Buffalo EA, Clark RE (2000) Impaired recognition memory in monkeys after damage limited to the hippocampal region. J Neurosci 20:451–463