Neurosci Bull October 1, 2013, 29(5): 553–564. http://www.neurosci.cn
DOI: 10.1007/s12264-013-1367-2

·Original Article·

What interests them in the pictures? — Differences in eye-tracking between rhesus monkeys and humans

Ying-Zhou Hu[1,2], Hui-Hui Jiang[1,2], Ci-Rong Liu[1,2], Jian-Hong Wang[1], Cheng-Yang Yu[1], Synnöve Carlson[3,4], Shang-Chuan Yang[1], Veli-Matti Saarinen[4], Joshua D Rizak[1,2], Xiao-Guang Tian[1], Hen Tan[1,2], Zhu-Yue Chen[1,2], Yuan-Ye Ma[1,5], Xin-Tian Hu[1]

[1]*State Key Laboratory of Brain and Cognitive Sciences and Key Laboratory of Animal Models and Human Disease Mechanisms, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China*

[2]*University of Chinese Academy of Sciences, Beijing 100049, China*

[3]*Institute of Biomedicine/Physiology, University of Helsinki, Helsinki, Finland*

[4]*Brain Research Unit, Aalto University School of Science, Espoo, Finland*

[5]*Kunming Primate Center, Chinese Academy of Sciences, Kunming 650223, China*

Corresponding authors: Jian-Hong Wang, Synnöve Carlson and Xin-Tian Hu. E-mail: wjh16@hotmail.com, syncarls@mappi.helsinki.fi, xthu@mail.kiz.ac.cn

ABSTRACT

Studies estimating eye movements have demonstrated that non-human primates have fixation patterns similar to humans at the first sight of a picture. In the current study, three sets of pictures containing monkeys, humans or both were presented to rhesus monkeys and humans. The eye movements on these pictures by the two species were recorded using a Tobii eye-tracking system. We found that monkeys paid more attention to the head and body in pictures containing monkeys, whereas both monkeys and humans paid more attention to the head in pictures containing humans. The humans always concentrated on the eyes and head in all the pictures, indicating the social role of facial cues in society. Although humans paid more attention to the hands than monkeys, both monkeys and humans were interested in the hands and what was being done with them in the pictures. This may suggest the importance and necessity of hands for survival. Finally, monkeys scored lower in eye-tracking when fixating on the pictures, as if they were less interested in looking at the screen than humans. The locations of fixation in monkeys may provide insight into the role of eye movements in an evolutionary context.

Keywords: eye-tracking; rhesus monkey; fixation; gaze

INTRODUCTION

Gaze direction, duration of fixation, and fixation count are indices used to investigate cognitive processes in animals, including rhesus monkeys and domestic horses[1-3]. With an eye-movement tracking system it is possible to study which features in a picture are attractive to subjects. Therefore, the comparison of eye-tracking data from human and non-human primates may help address the evolutionary features of perception and cognition between species, thus providing insights into the evolution of the cognitive processes in the brain.

When humans are shown a picture containing animal or human features, most fixate first on the faces in the picture. This is attributed to the importance of face recognition in the social interactions of humans[4]. Furthermore, the human brain has developed a complex cognitive system of eye fixation on the perceptual elements

of faces and eyes, presumably because eyes are a more attractive feature than the other parts of the face.

The eye region containing eyes, eyebrows, eyelids and eyelashes, plays a key role in face recognition because it is the most attended to among all facial features and the most-used source of information during social interactions[5-10]. Besides, the scanning of a face always starts from the eyes[11].

Some studies investigating the pattern of eye-movements in monkeys have demonstrated that they, like humans, fixate intensely on the eye region of faces when viewing conspecific face images[12, 13]. However, very few studies have used both humans and monkeys viewing naturalistic pictures that contain full-body images of animals including humans. In one such study in which the eye-movements of chimpanzees were recorded[14], the chimpanzees were strikingly similar to humans in the viewing patterns: they both looked for longer periods at animal figures than at the background, and at the face region longer than at other parts of the body. However, the two species showed differences in their fixation shift between areas of interest (AOIs) and fixation duration on the face region. The AOIs in our experiment were the foreground objects, humans and animals, based on many studies on cultural differences between humans[15] and animal eye-tracking[14].

It is not surprising that the viewing patterns of chimpanzees and humans are similar due to the close evolutionary ties between the species. However, different non-human primate species have different evolutionary relationships with humans. In addition to the importance of face recognition in social interactions, the hands and food handling are also important for the evolution and survival of both humans and monkeys[16, 17]. In the current study, we investigated eye-tracking in pictures containing either monkeys, humans or both, with specific interest in whether there are differences between these two species in how they look at the eyes, face, hands, and what is being done with the hands.

SUBJECTS AND METHODS

Human Participants and Animals

Six male rhesus monkeys (*Macaca mulatta*) from the breeding colonies at Kunming Institute of Zoology (KIZ) were used. Their mean age was 6.5 ± 0.7 years and mean body weight 8.5 ± 0.9 kg. The monkeys were housed singly under standard conditions (under a 12-h light/dark cycle with the light on from 07:00 to 19:00; humidity at 60%, 21 ± 2°C) in the animal house.

The experiments were conducted in accordance with the guidelines for the National Care and Use of Animals and approved by the Chinese National Animal Research Authority.

Twenty healthy undergraduate and graduate students (10 males and 10 females; mean age, 26.4 ± 0.5 years) from KIZ participated in the eye-tracking study. They were paid 10 RMB in compensation.

Apparatus

Eye-tracking was performed with a Tobii X120 Studio Eye Tracker (Tobii Technology AB, Stockholm, Sweden) mounted on a table in a separate testing room which could be isolated by a door.

Human subjects were seated in front of the camera in the testing room. Monkeys were seated in a monkey chair in front of the camera. A 22-inch LCD monitor[18] (SAMSUNG 2233RZ) with a resolution of 1280×1024 was placed behind and above the camera. The distance between the monitor and the camera was 41 mm. The distance between the camera and the subjects' eyes was 67.1 ± 0.8 cm. The active display area of the monitor was 370 mm in width and 295 mm in height and all pictures were shown within this area. The physical dimensions of the pictures were not identical. The viewing angle did not exceed 35° to any point on the screen.

The eye-tracking was conducted in the testing room with the door closed, while the experimenter collected the data in the adjacent room. Both the humans and monkeys participated in the experiment in the same way.

Among the pictures presented to human subjects and monkeys, part of them are from International Affective Picture System (IAPS), authorized by NIMH Center For the Study of Emotion and Attention, and the rest are taken by the authors, especially the pictures containing the primates.

Procedure

Calibrations were conducted before each test using a regular 5-point calibration with a medium speed for all subjects. Audio cartoon symbols were also used as

calibration points for some monkeys less interested in the screen, to attract their attention. For human subjects, the calibration was conducted once if both eyes were successfully captured. For monkeys, less than half of the calibrations were successful in one eye for all dots. Therefore, repeated calibrations were conducted. Usually, the calibration was done three times in each monkey.

In the experiment, two sets of 30 pictures each were presented. The order of presentation was randomized between subjects. Each set contained 11 monkey pictures, seven human pictures, nine mixed human and animal pictures (one picture contained a dog and a child, and eight contained monkeys and humans) and three cartoon face pictures (presented randomly among the human and monkey pictures). The three cartoons were faces with eyes looking to the front, left and right. Each picture was shown for 5 s, with an interstimulus interval (ISI) of 2 s. During the ISI, a white fixation cross ('X') was presented at the center bottom of the screen.

Monkeys could move their heads ~1 cm (~1° visual angle) forward and backward due to the collar used for holding the monkey in the primate chair (diameter ~1 cm larger than the monkey's neck). The excess 1 cm was used to keep the monkey comfortable, while preventing any larger movements. Besides, the monkeys could still turn the body or head. Such turning led to a failure of capture by the Tobii studio system. Unsuccessful captures were reflected by a zero in the recording and were excluded from analysis.

Each monkey was tested on different days, the tests being carried out at least four times. Each test lasted ~40 min and contained the two sets of pictures. The monkeys were led to the monkey chair in the testing room at least 30 min prior to the start of the experiment for habituation.

Monkeys received a piece of fruit after the test. If the monkey was impatient and did not continue watching the monitor, an experimenter would go into the testing room and accompany the monkey to make it feel comfortable. If the monkey did not look at the screen, the experimenter then gently fixed the monkeys' head by using a padded wooden box to hinder the monkey's ability to turn its head.

Each human participant was instructed to sit without moving his/her head or sitting position during the calibration and the test of each set of pictures. They were asked to direct their gaze at the fixation cross between the images and to freely look at the pictures when they appeared. Each participant was tested only once with each of the two sets of pictures. Between the two tests, participants were allowed to rest.

Data Collection

Several AOIs were defined within each picture: the eyes, the head (whole head excluding the eyes), the body (whole body excluding the head and hands), and hand + handling (including the hands and objects handled). The frames of the eyes, head and body were drawn close to their actual outlines. There was no overlap between any two frames (see Fig. 1 for example).

The data collection was started 200 ms after the beginning of picture presentation so the subjects had time to move their eyes to the area of the picture, and lasted for 4.8 s.

Fig. 1. Examples of monkey (A), human (B), and animal + human (C) pictures presented in this study. Areas of interest are marked in different colors, eyes in red, head in yellow, body in green, and hand + handling in purple. Eye-A means eyes of the animal, and Body-H means body of the human, *etc.*

The minimum fixation duration was set at 75 ms. A fixation event <75 ms was considered to be a saccade.

Each monkey was tested more than four times for each picture because it could not concentrate on all the pictures during the whole test. We then first calculated the mean score for each monkey for all successful captures. For each AOI, the fixation length (total fixation duration), fixation count, and the first fixation duration were calculated. Normalization of fixation length was done as follows: normalized fixation length = percentage of fixation length/percentage of AOI, in which percentage of fixation length = fixation length in each AOI/ fixation length in total in picture×100%; percentage of AOI = area of each AOI/area of total picture × 100%. Normalization of the fixation count was calculated in the same way. First fixation duration was not normalized.

In addition, the percentage of participants fixating on each AOI was calculated.

The fixation results from the three sets of pictures were separately analyzed between groups.

Data Analysis

Statistical AOI analysis was performed using SPSS 13.0. Differences in scores between groups were assessed by analysis of variance (ANOVA) with repeated measures where appropriate. Subject (monkeys and humans) was considered as a between-group factor, whereas the AOIs of the picture (eye, head, body and hand+handling) and not-AOI (fixation areas outside the AOIs) were considered as within-group factors. Greenhouse-Geisser correction was used when sphericity did not hold. Bonferroni confidence interval adjustment was used when pair-wise comparisons were conducted for AOIs within each group. One-way ANOVA was used to further analyze the difference between groups in their fixation AOIs. Differences were considered significant when $P \leq 0.05$.

RESULTS

Pictures Containing Monkey Features

Fixation length The fixation length was measured as the time spent focusing on one of the AOIs in the pictures. There were main effects for group [$F(1,24) = 21.24$, $P < 0.001$] and AOI [$F(4,96) = 59.95$, $P < 0.001$] as well as an interaction between group and AOI [$F(4,96) = 96.00$, $P < 0.001$].

A one-way ANOVA showed that monkeys paid less attention to and spent less time than humans fixating on the head, and hand+handling features in pictures of monkeys (Fig. 2A, Table 1), but they had body fixation similar to the human participants. Meanwhile, they had longer fixation lengths on the background area in these pictures than the humans (Table 1). The monkeys also tended to look less on the eye than the human subjects in pictures of monkeys, though the tendency did not reach a significant level (discussed later) (Fig. 2A, Table 1).

The group analyses of the fixation length on the AOIs in the monkey pictures showed that monkeys spent the longest time on the hand+handling area, then the

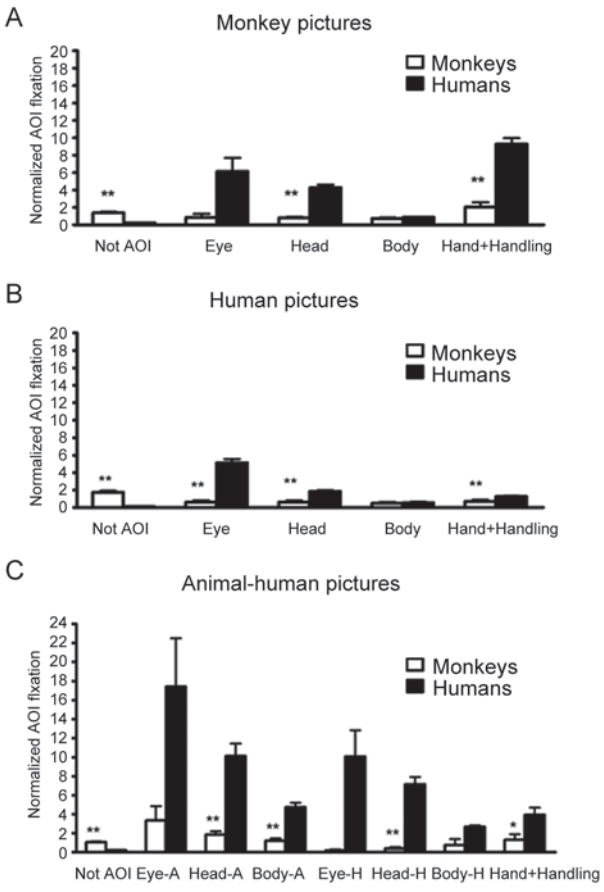

Fig. 2. Normalized fixation length within 4.8 s in each AOI of each set of pictures for monkeys and humans. "Not AOI" indicates parts of the picture other than the AOIs. Eye-A means eyes of the animal, and Body-H means body of the human, *etc.* *P <0.05, **P <0.01.

Table 1. Statistical results for monkey and human fixation on pictures containing monkey features

	Not AOI	Eye	Head	Body	Hand+Handling
Fixation length					
F(1,25)	204.97		27.38		30.34
P value	<0.001	n.s.	<0.001	n.s.	<0.001
Fixation count					
F(1,25)	195.80		37.10	7.95	63.15
P value	<0.001	n.s.	<0.001	0.009	<0.001
First fixation duration					
F(1,25)	8.89	4.28	24.81	8.65	15.06
P value	0.006	0.049	<0.001	0.007	0.001
Percentage of participants fixating at the AOI					
F(1,25)		15.09	73.09	100.16	112.36
P value	n.s.	0.001	<0.001	<0.001	<0.001

n.s., not significant.

background [$F(4,33) = 3.03$, $P = 0.003$]. They gazed at the other parts of the pictures for shorter times than hand+handing [$F(4,33) = 3.03$, $P = 0.029$ for eye, 0.012 for head, and 0.008 for body]. Humans, on the other hand, were more attracted by the hand+handing area than the eye [$F(4,95) = 21.9$, $P = 0.003$], head [$F(4,95) = 21.9$, $P < 0.001$], body [$F(4,95) = 21.9$, $P < 0.001$], or not-AOI [$F(4,95) = 21.9$, $P < 0.001$]. It was also noteworthy that humans concentrated more on the eye than the other parts except for hand+handing in the monkey pictures (Fig. 2A).

Fixation counts Fixation counts were measured as the number of times a subject fixated on an AOI. Interestingly, similar results were found after normalization (Fig. 3A), suggesting that humans and monkeys had different fixation interests in the picture features [main effects of group, $F(1,24) = 96.00$, $P < 0.001$; AOI, $F(4,96) = 99.11$, $P < 0.001$; interaction between AOI and group, $F(4,96) = 100.40$, $P < 0.001$]. Monkeys had lower fixation counts than humans at all AOIs in the monkey pictures. But the opposite occurred for the background of the pictures (Table 1). Furthermore, monkeys fixated more often on the hand+handling than the background [$F(4,33) = 1.76$, $P = 0.041$]. The humans also fixated more often on the hand+handling *versus* the other AOIs or background [$F(4,95) = 53.14$, all $P < 0.001$], and on the eye and head *versus* the body or background [$F(4,95) = 53.14$, all $P < 0.001$] (Fig. 3A).

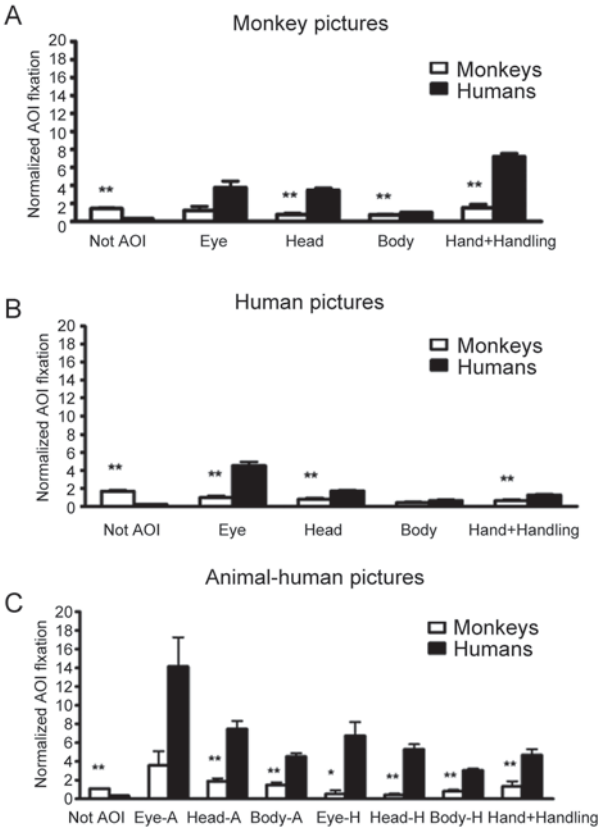

Fig. 3. Normalized fixation count within 4.8 s in each AOI of each set of pictures in monkeys and humans. Eye-A means eyes of the animal, and Body-H means body of the human, *etc.* **P* <0.05, ***P* <0.01.

First fixation duration The first fixation duration was defined as the amount of time the subject first spent looking at an AOI, excluding the first 200 ms, before changing the direction of attention.

Different from the fixation length and fixation count results, the first fixation duration on the monkey pictures showed that subjects had different interests in the AOIs depending on the group [main effect of group on AOI, $F(4,96) = 11.37$, $P < 0.001$], as well as the interaction between AOI and group [$F(4,96) = 23.45$, $P < 0.001$]. In detail, the monkeys' first fixation duration at certain AOIs (body, head, hand+handling and eye) was much shorter than the humans' first glance. And the opposite occurred for not-AOI (Fig. 4A, Table 1). Furthermore, monkeys had a longer first glance at the body than the eye [$F(4,33) = 9.49$, $P < 0.001$] and hand+handling [$F(4,33) = 9.49$, $P = 0.015$]. The humans gazed at first sight longer at the head *versus* the eye, body, or not-AOI [$F(4,95) = 15.3$, all $P < 0.001$], and at hand+handing *versus* the eye [$F(4,95) = 15.3$, $P = 0.007$], body [$F(4,95) = 15.3$, $P = 0.011$], or not-AOI [$F(4,95) = 15.3$, $P < 0.001$] (Fig. 4A).

Percentage of participants fixating on the AOIs A lower percentage of monkeys gazed at the eyes, head, body and hand+handling compared than humans at a corresponding AOI in the monkey pictures (Fig. 5A). No significant difference between the two groups was found for the background in the pictures (Table 1).

Pictures Containing Human Features

Fixation length Significant differences were found

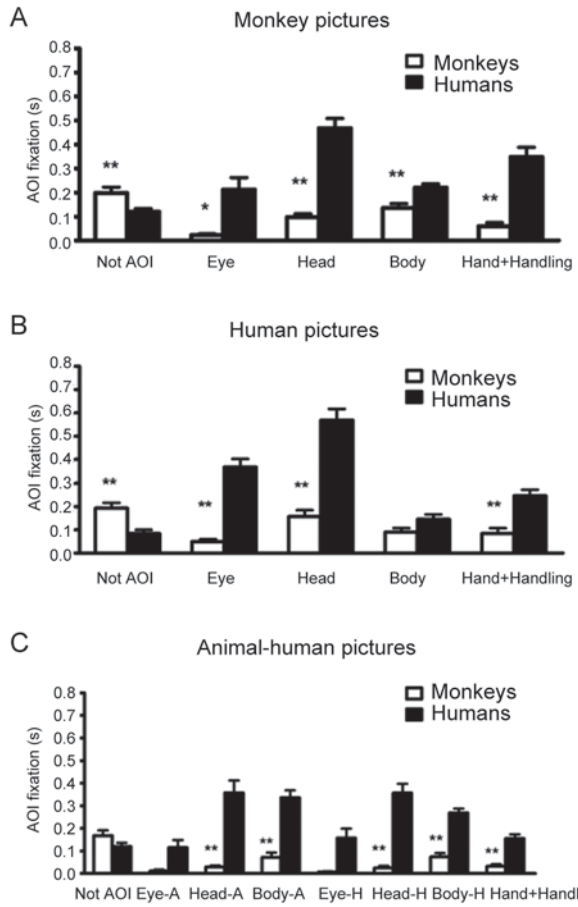

Fig. 4. First fixation duration within 4.8 s on each AOI of each set of pictures in monkeys and humans. Eye-A means eyes of the animal, and Body-H means body of the human, *etc.* *$P < 0.05$, **$P < 0.01$.

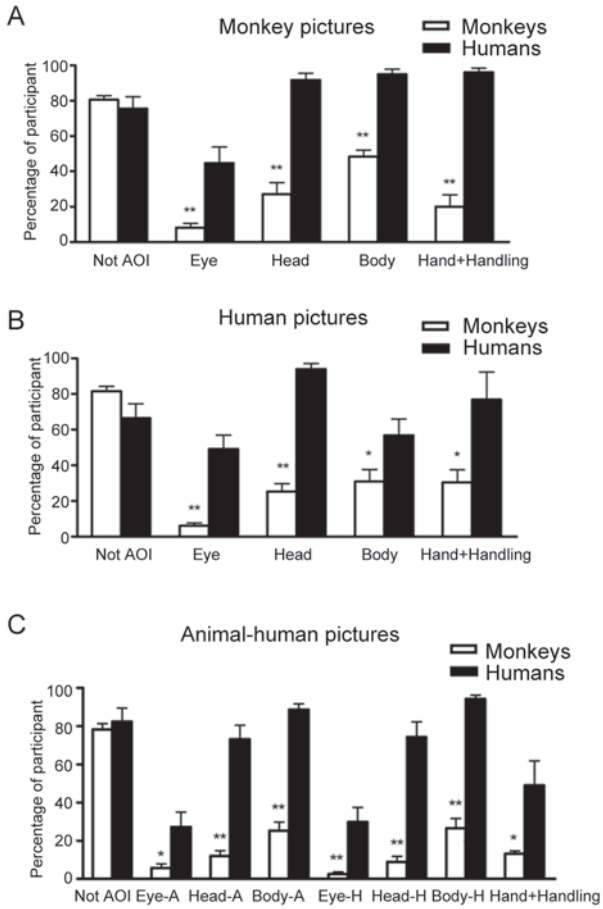

Fig. 5. Percentage of participants fixating on each AOI of each set of pictures in monkeys and humans. Eye-A means eyes of the animal, and Body-H means body of the human, *etc.* *$P < 0.05$, **$P < 0.01$.

between the two groups in the fixation length on the AOIs in pictures containing human features [main effect of AOI, $F(4,96) = 38.41$, $P < 0.001$] (Fig. 2B). These differences depended on the group [interaction between AOI and group, $F(4,96) = 94.44$, $P < 0.001$]. Specifically, monkeys fixated for a shorter time than humans on the eye, head, and hand+handling in the human pictures (Fig. 2B, Table 2). Further, monkeys had a shorter fixation length on the AOIs than on the background [$F(4,25) = 11.07$, all $P < 0.001$], and they spent almost equal time on the human's eye, head, body, and hand+handling (Fig. 2B). In contrast, the human subjects were more interested in the human eye among these features than others [$F(4,95) = 92.47$, all $P < 0.001$].

Fixation counts The fixation counts for the human images differed between the AOIs within subjects [main effect of AOI, $F(4,96) = 44.00$, $P < 0.001$] and this difference depended on the group [AOI × group: $F(4,96) = 70.05$, $P < 0.001$] (Fig. 3B).

Monkeys had a lower fixation count than humans for the eye, head, and hand+handling but not for the body area, and a higher fixation count for the area outside the AOIs in the pictures (Table 2). However, humans gazed at the eye more often than at the head, hand+handling, body, and not-AOI [$F(4,95) = 68.53$, all $P < 0.001$] in these pictures (Fig. 3B).

First fixation duration The duration of the first fixation differed between the groups and between the AOIs [main effect of group: $F(1,24) = 9.23$, $P = 0.006$; main effect of AOI: $F(4, 96) = 19.64$, $P < 0.001$; interactive effect AOI × group: $F(4,96) = 23.43$, $P < 0.001$] (Fig. 4B).

The monkeys had a longer first fixation duration on the not-AOI, and a shorter duration on the eye, head or hand+handling of the human images compared with the human participants (Table 2). Further, monkeys also tended to gaze longer at the head rather than the eye [$F(4.25) = 5.45$, $P = 0.001$] at first sight. Similarly, humans spent a longer time looking at the head on first glance, than at the eye, body, hand+handling and not-AOI [$F(4,95) = 38.5$, all $P < 0.001$]. In humans, the duration of the first fixation was longer for hand+handling than for the background [$F(4,95) = 38.5$, $P < 0.001$], and longer for the eye than not-AOI [$F(4,95) = 38.5$, $P < 0.001$], body [$F(4,95) = 38.5$, $P < 0.001$] and hand + handling [$F(4,95) = 38.5$, $P = 0.007$] when they first fixated on the hand+handling or eye portions of the pictures.

Percentage of participants fixating on the AOI Fewer monkeys than humans were found to gaze at the eye, head, body and hand+handling (Fig. 5B, Table 2). However, there was no difference in the percentage of participants fixating on not-AOI.

Pictures Containing Both Human and Primate Features

Table 2. Statistical results of the comparisons between monkey and human fixation on pictures containing human features

	Not-AOI	Eye	Head	Body	Hand+Handling
Fixation length					
$F(1,25)$	210.42	34.07	27.56		15.83
P value	<0.001	<0.001	<0.001	n.s.	0.001
Fixation counts					
$F(1,25)$	259.08	20.76	23.41		15.07
P value	<0.001	<0.001	<0.001	n.s.	0.001
First fixation duration					
$F(1,25)$	11.55	23.57	21.15		11.56
P value	0.002	<0.001	<0.001	n.s.	0.002
Percentage of participants fixating on the AOI					
$F(1,25)$		29.09	170.51	5.16	7.49
P value	n.s.	<0.001	<0.001	0.042	0.034

n.s., not significant.

Fixation length The fixation length of monkeys and humans on pictures containing both human and animal features is depicted in Fig. 2C. A main effect was found for the AOI [$F(7,168) = 132.92$, $P < 0.001$], and for the interaction between AOI and group [AOI × group, $F(7,168) = 80.51$, $P < 0.001$]. Since each picture contained features of both animals and humans, we analyzed the AOIs in two sets: the eye, head, and body AOIs of animals (A) and humans (H). The hand+handling AOIs for both humans and animals in a picture were treated as one AOI, due to their small areas. Monkeys fixated for a shorter time than humans on all AOIs (Table 3). Further, monkeys looked longer at animals' eyes than at humans' eyes [$F(7,40) = 2.33$, $P = 0.003$], head [$F(7,40) = 2.33$, $P = 0.005$], and hand+handing [$F(7,40) = 2.33$, $P = 0.017$]. Furthermore, the monkeys fixated longer than humans on the background of the pictures, as they had done with the pictures containing only primates or humans (Fig. 2C).

Human participants spent longer gazing at the animals' eyes than the head [$F(7,152) = 5.39$, $P = 0.037$], body [$F(7,152) = 5.39$, $P < 0.001$], and hand+handling [$F(7,152) = 5.39$, $P < 0.001$]. Similar results were found for humans gazing at human features (eye > head > hand+handling > body).

Fixation counts The fixation counts for the pictures containing both species differed between AOIs within the groups [main effect of AOI, $F(7,168) = 178.00$, $P < 0.001$] and also depended on the group [AOI × group, $F(7,168) = 81.58$, $P < 0.001$] (Fig. 3C).

The monkeys gazed less than humans at both sources of eye, head, body and hand+handling AOIs. Moreover, the monkeys visited the areas outside the AOIs more often than the humans (Table 3).

First fixation Duration There were significant differences in the first fixation duration between the AOIs within subjects [main effect of AOI, $F(7,168) = 16.30$, $P < 0.001$] (Fig. 4C).

The monkeys had a shorter first fixation duration than humans on both kinds of head and body, and hand+handling (Table 3). They also fixated for less time on the eye than the human subjects, though the difference did not reach a significant level (discussed later) (Fig. 4C, Table 3). Moreover, the monkeys had a longer first fixation on body-A than eye-A [$F(7,40) = 14.35$, $P = 0.004$], as well as on body-H than eye-H [$F(7,40) = 14.35$, $P = 0.001$] or hand+handling [$F(7,40) = 14.35$, $P = 0.039$]. Besides, the humans had a longer first fixation on head-A than on eye-A, eye-H and not-AOI, as well as on head-H than eye-A,

Table 3. Statistical results for monkey and human fixation on pictures containing both human and monkey features

	Not AOI	Eye-A	Head-A	Body-A	Eye-H	Head-H	Body-H	Hand+Handling
Fixation length								
F(1,25)	220.01		10.55	18.70		20.42		5.03
P value	<0.001	n.s.	0.003	<0.001	n.s.	<0.001	n.s.	0.034
Fixation counts								
F(1,25)	159.23		13.69	22.50	5.67	23.00	45.77	8.17
P value	<0.001	n.s	0.001	<0.001	0.026	<0.001	<0.001	0.009
First fixation duration								
F(1,25)			10.41	19.77		19.47	28.89	12.83
P value	n.s.	n.s.	0.004	<0.001	n.s	<0.001	<0.001	0.002
Percentage of participants fixating on the AOI								
F(1, 25)		7.27	62.10	136.36	12.24	59.95	152.84	7.69
P value	n.s.	0.022	<0.001	<0.001	0.006	<0.001	<0.001	0.024

n.s., not significant. Eye-A, Head-A, and Body-A mean fixation on animal eye, head, and body, respectively. Eye-H, Head-H, and Body-H mean fixation on the corresponding part of human.

eye-H, hand+handling and not-AOI [$F(7,152) = 9.83$, $P < 0.001$ for all] (Fig. 4C).

Percentage of participants fixating on the AOI The percentage of humans fixating on the eye, head, body or hand+handling was higher than the corresponding percentage of monkeys. However, there was no difference in the percentage of participants fixating on not-AOI between monkeys and humans (Fig. 5C, Table 3).

DISCUSSION

In this study, we demonstrated that humans always preferred looking at the eyes in all the pictures and were interested in what was done with the hands in monkey pictures as well as in pictures containing both monkeys and humans. But monkeys seemed to look randomly at both human and monkey pictures. For pictures containing both humans and animals (mostly monkeys), the monkeys seemed to prefer looking at the eyes, head and body of the animals and avoid looking at the eyes of humans (Figs. 2 and 3). They also avoided looking at the eyes in all three sets of pictures at the first sight (Fig. 4).

The Three Sets of Pictures

In terms of normalized fixation lengths and fixation counts, when gazing at pictures containing only non-human primates or humans, most interestingly, rhesus monkeys avoided looking at the eyes in all the three types of pictures, including at first sight (Fig. 4). These findings are in line with observations in monkey societies in their natural environment. Monkeys usually avoid looking straight at each other's eyes as the gaze might signal an attack or foster aggressive behavior[19, 20]. This result is also partly consistent with other primate studies[14]. The human participants, on the other hand, were more interested in the animal or human eyes and head than in the other parts of the images. This finding is consistent with the social character of humans, who communicate directly with each other, often by receiving feedback *via* other people's responses including their facial expressions[21, 22]. In the current study, the eyes and the head were efficient areas for subjects to receive useful information. We believe this interest in the eyes and head area can be attributed to the social nature of humans.

Although the monkeys paid less attention to the AOIs than the humans did, they still had a pattern similar to the humans in gazing at what was being done with the hands, especially in the pictures containing monkeys (Figs. 2A and 3A). Their interest in the hand+handling area tended to be higher than that in not-AOI (Fig. 2A). This might be attributed to the monkey's attention to the food or plant that was being held in the pictures.

When the humans first looked at the hand area of the pictures, they inspected the hands and the objects being held for a long time. In addition, the percentage of humans gazing at the area with a hand handling objects was similar (almost identical) to the percentage of those looking at an animal's head or body. This suggests that humans and monkeys were curious about what was being done with the animal's hands, in addition to the other important information in the picture.

When subjects looked at the pictures containing humans or monkeys, the monkey's first fixation duration was longest for the body area in monkey pictures whereas in human pictures it was for the head, among areas other than not-AOI. The first fixation duration of the humans, on the other hand, was longest for the head and then the hand holding objects in the monkey pictures whereas in the human pictures it was the head and then the eye area. This suggests that (1) monkeys are more curious or may be more conditioned to the monkey body than to the head and eyes, but more curious about the human head than body. This may be explained by reduction of the monkey's aversion to the human head, as this is a less threatening or less aggressive signal because it is not of the same species, which may also have increased the monkeys' curiosity; and (2) the humans are most interested in the human head and eyes, which may be due to the socially valuable information they deliver. A higher percentage of humans looked at the hands and what they were handling in the monkey pictures than in the human pictures. This may have occurred because the humans were not familiar with what was being done with the monkey's hands.

With respect to pictures containing both monkey and human features, monkeys looked at the animal's eye, head and body for a longer time than at the humans' features (Fig. 2C). This result is inconsistent with the result from the pictures containing only monkeys or humans, which suggests that the combined pictures might be more efficacious for testing the interests of the animals than pictures only containing one species.

We demonstrated that the results from humans were distinct from those of monkeys (Figs. 2A, 2C, 3A and 4C); however, the statistical analysis showed a non-significant difference due to the small area and variation of eyes (range from 0.09% to 8.1% of the whole picture), which might have made it difficult to notice the small eyes. Another reason is the relatively small samples in the groups, requiring a further study that involves more subjects and bigger eyes in the pictures.

On the other hand, it is important to note that eye-tracking may only reflect a subject's interest and/or attention up to a certain point. In some species including rhesus monkeys, eye-tracking does not reveal their interest and attention as reliably as in humans, because monkeys may pay attention to one another without directly looking at each other[19, 20]. This may also provide another rationale for why the monkeys tested in this study failed to show curiosity about the monkey's head in the pictures but did show curiosity about the human's head.

At first glance, monkeys watched the body for a long period, whereas humans showed a high interest in the head and the body of the two species. This result was partly consistent with the findings with the other two sets of pictures, suggesting that monkeys might prefer looking at the body in a picture containing the image of a creature in order to avoid seeing their face and eyes. However, there was a difference in the results between the first fixation duration time and fixation lengths/counts, as the former data were not normalized. Since the first fixation duration might reflect the instinctive interest of the subjects, it was inappropriate to normalize the raw data.

It is possible that the monkeys did not spend a long time looking at the human's head in pictures containing both animals and humans because the AOIs were smaller in this set of pictures than in those containing humans only. Similarly, the human subjects did not pay much attention to the eyes at first sight since the area of the eyes was small, thus not providing enough useful information to keep their attention. But humans were still more curious about the details of the face during their first fixation.

Different Indices

Similar exploratory behavior has been analyzed in mice using a Y-maze. According to Dellu *et al.*[23, 24], two separate measures (inspective and inquisitive behaviors) can be analyzed in the Y-maze. The inspective exploratory behavior is reflected by the time spent in the Y-maze arms, and the number of arm visits is an index of inquisitive behavior. In the current study, we used a similar concept, assuming that the fixation length on each AOI might be an index of the subject's inspective behavior towards a particular AOI, and the fixation count might reflect the inquisitive behavior or the subject's curiosity toward each AOI. Furthermore, the first fixation duration was considered an indication of how much attention the subjects paid to the AOI, and the percentage of participants that looked at an AOI a measure of the general curiosity of each group.

With respect to the first fixation duration, which can be used as an index of the attention that the subjects pay to an AOI, the monkeys attend to the body in monkey pictures, the head in human pictures, and the head and body in pictures containing both monkeys and humans. However, at first glance, the humans paid more attention to the head and eyes than to the body, as measured by the duration of the first fixation, suggesting that humans are interested in things related to their social communicative needs.

As for the general curiosity in each group, reflected by the percentage of participants that looked at an AOI during the eye-tracking, monkeys tended to be interested in the body, the head and then the hand handling objects. Humans were interested in the head, eyes, body, and then the hand handling objects. However, it is important to note that a lower percentage of monkeys than humans looked at the AOIs, suggesting that the monkeys were not as interested in fixating on the details in the pictures as the humans. Furthermore, the percentages of monkeys and humans viewing the background were similar. Therefore, it can be inferred that most of the monkeys watched the pictures throughout the eye-tracking, although they did not pay as much attention as the humans to the different AOIs in the pictures. Instead, the monkeys inspected other areas of the pictures more often. One possibility is that they did not grasp the meaning of the pictures, whereas the humans understood the framework and were able to focus on the details.

In addition to the main features (AOIs) that the animals and humans attended to, the use of eye-tracking to measure the interest in what is being done with the hand

(hand+handling) is an important index, because it provides information essential for learning, mimicry and survival[16, 17].

In the present experiment, the monkeys were all allowed to become accustomed to being seated in the primate chair before the experiment started in order to reduce stress. Although they were not trained to fixate on the screen beforehand, four out of the six looked at the pictures very well and most seemed attracted by the pictures. Nonetheless, we tested each monkey's gaze at the pictures at least four times. The current study estimated the monkeys' primary instinct when they were required to gaze at pictures containing the whole body of monkey, human or both, presented in an LCD monitor. The eye movements of the monkeys were monitored and recorded using a eye-tracker. Therefore, the paradigm used may be close to reflecting the monkey's natural inclination.

However, monkeys showed lower concentration on the pictures than humans and a random pattern of fixation when assessed by normalized fixation length and normalized fixation count. An explanation for the lower concentration might be that the monkeys were not trained to fixate on the screen before the experiment. Another possibility is that the monkeys were less interested in fixating on items and more interested in piecing together the whole picture, which may cause a random fixation pattern during the process. Further work is needed to compare the differences in fixation between trained monkeys and human subjects and to elucidate why the monkeys displayed what appeared to be lower concentration.

The comparison of gaze between monkeys and humans may help to address certain evolutionary features of perception and cognition in animals and humans, and provide insights into some neurological disorders. For example, impaired or abnormal eye movements, gaze or fixation may indicate certain disorders such as autism[25-27]. Therefore, it is necessary to further investigate and compare gaze and fixation in trained monkeys, healthy human subjects and humans with mental disorders.

ACKNOWLEDGMENTS

We greatly appreciate Dr. Riitta Hari for suggestions on the manuscript. This work was supported by the National Basic Research Development Program (973 program) of China (2012CB825500, 2011CB707800), Basic Research Frontier Project of Chinese Academy of Sciences, China (KSCX2-EW-J-23), the National Natural Science Foundation of China (31271167, 31271168, 81271495, 31070963, 31070965), the Strategic Priority Research Program of the Chinese Academy of Sciences, China (XDB02020000), the Academy of Finland (National Centers of Excellence Program 2006-2011, Grant No. 259752) and the aivoAALTO Project of the Aalto University, Finland.

Received date: 2012-09-18; Accepted date: 2013-03-12

REFERENCES

[1] Flombaum JI, Junge JA, Hauser MD. Rhesus monkeys (*Macaca mulatta*) spontaneously compute addition operations over large numbers. Cognition 2005, 97: 315–325.

[2] Proops L, McComb K, Reby D. Cross-modal individual recognition in domestic horses (*Equus caballus*). Proc Natl Acad Sci U S A 2009, 106: 947–951.

[3] Sulkowski GM, Hauser MD. Can rhesus monkeys spontaneously subtract? Cognition 2001, 79: 239–262.

[4] Hari R, Kujala MV. Brain basis of human social interaction: from concepts to brain imaging. Physiol Rew 2009, 89: 453–479.

[5] Henderson JM, Williams CC, Falk RJ. Eye movements are functional during face learning. Mem Cogn 2005, 33: 98–106.

[6] Janik S, Wellens A, Goldberg M, Dell'Osso L. Eyes as the center of focus in the visual examination of human faces. Percept Mot Skills 1978, 47: 857–858.

[7] Laughery KR, Alexander JF, Lane AB. Recognition of human faces: Effects of target exposure time, target position, pose position, and type of photograph. J Appl Psychol 1971, 55: 477–483.

[8] Luria SM, Strauss MS. Comparison of eye movements over faces in photographic positives and negatives. Perception 1978, 7: 349–358.

[9] Itier RJ, Villate C, Ryan JD. Eyes always attract attention but gaze orienting is task-dependent: evidence from eye movement monitoring. Neuropsychologia 2007, 45: 1019–1028.

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

[11] Schyns PG, Petro LS, Smith ML. Dynamics of visual information integration in the brain for categorizing facial expressions. Curr Biol 2007, 17: 1580–1585.

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

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

[14] Kano F, Tomonaga M. How chimpanzees look at pictures:

a comparative eye-tracking study. Proc Biol Sci 2009, 276: 1949–1955.

[15] Goh JO, Tan JC, Park DC. Culture modulates eye-movements to visual novelty. PLoS One 2009, 4: e8238.

[16] Pouydebat E, Reghem E, Borel A, Gorce P. Diversity of grip in adults and young humans and chimpanzees (*Pan troglodytes*). Behav Brain Res 2011, 218: 21–28.

[17] Hopkins WD, Leavens DA. Hand use and gestural communication in chimpanzees (*Pan troglodytes*). J Comp Psychol 1998, 112: 95–99.

[18] Wang P, Nikolić D. An LCD monitor with sufficiently precise timing for research in vision. Front Hum Neurosci 2011, 5.

[19] Thomsen CE. Eye contact by non-human primates toward a human observer. Anim Behav 1974, 22: 144–149.

[20] Mendelson MJ, Haith MM, Goldmanrakic PS. Face scanning and responsiveness to social cues in infant rhesus monkeys. Dev Psychol 1982, 18: 222–228.

[21] Trevarthen C. Facial expressions of emotion in mother-infant interaction. Hum Neurobiol 1985, 4: 21–32.

[22] Dimberg U, Thunberg M, Elmehed K. Unconscious facial reactions to emotional facial expressions. Psychol Sci 2000, 11: 86–89.

[23] Dellu F, Contarino A, Simon H, Koob GF, Gold LH. Genetic differences in response to novelty and spatial memory using a two-trial recognition task in mice. Neurobiol Learn Mem 2000, 73: 31–48.

[24] Dellu F, Mayo W, Cherkaoui J, Le Moal M, Simon H. A two-trial memory task with automated recording: study in young and aged rats. Brain Res 1992, 588: 132–139.

[25] Takarae Y, Minshew NJ, Luna B, Krisky CM, Sweeney JA. Pursuit eye movement deficits in autism. Brain 2004, 127: 2584–2594.

[26] Kemner C, Verbaten MN, Cuperus JM, Camfferman G, van Engeland H. Abnormal saccadic eye movements in autistic children. J Autism Dev Disord 1998, 28: 61–67.

[27] Pensiero S, Fabbro F, Michieletto P, Accardo A, Brambilla P. Saccadic characteristics in autistic children. Funct Neurol 2009, 24: 153–158.