Face scanning in chimpanzees and humans: Continuity and discontinuity

Research Article

Running title: Face scanning in chimpanzees and humans

Word count: 6584 words

This article includes a table, three figures, and five supplementary electronic addenda (four figures and a video)

Abstract

How do chimpanzees, the species with the closest evolutionary connection to humans, view faces? This study is the first to use the eye-tracking method to perform direct comparisons between humans and chimpanzees with regard to face scanning. Members of both species viewed the same sets of photographs representing conspecific and non-conspecific faces under the same experimental conditions. Chimpanzees and humans exhibited systematic and similar patterns of face scanning, including intensely viewing main facial features (i.e., eyes, nose, and mouth) and inspecting the eyes and mouth, in that order. However, several differences between the species were also evident. For example, humans were more likely to exhibit sequential re-fixations on the eye regions than were chimpanzees, whereas chimpanzees were more likely to engage in quick, vertical scanning over the eyes and mouth. Such species similarities and differences were consistent across conspecific and non-conspecific faces and were thus independent of the external morphologies of species-specific faces. Furthermore, when presented with facial expressions, chimpanzees changed their scanning patterns in response to those facial actions, whereas humans maintained intense eye-viewing across the expressions. Finally, we discuss how these face scanning patterns are related to species-specific forms of facial communications in chimpanzees and humans, and suggest that both species have unique eye movement strategies for interactions with conspecifics.

Key words: chimpanzees, emotional expressions, eye-tracking, evolution, faces.

Introduction

Humans have highly sophisticated forms of facial communication. Faces and eye movements contribute to a variety of expressions that are then efficiently perceived in humans. Moreover, humans often engage in lengthy face-to-face communications, accompanied by intense eye contact (Argyle & Cook, 1976). How and when such unique forms of human communications evolved have long been of interest, particularly from a comparative perspective that addresses how and to what extent human and non-human animals are similar and dissimilar in their forms of communications (Chevalier-Skolnikoff, 1973; Darwin, 1999). Chimpanzees (*Pan troglodytes*), the species closest to humans, are known to share several forms of facial communications with humans. Field-observational, morphological, and experimental studies have identified similarities between the species. For example, both frequently engage in mutual gazing during mother-offspring interactions (Bard et al., 2005; van Lawick-Goodall, 1967). Additionally, the facial musculature of chimpanzees is homologous with that of humans in several respects, and both species exhibit some similar facial expressions (e.g., play faces: Chevalier-Skolnikoff, 1973; Parr, Waller, Vick, & Bard, 2007; Vick, Waller, Parr, Pasqualini, & Bard, 2007). Experimental studies have shown that the mechanisms underpinning face/gaze perception in chimpanzees are similar to those in humans, including the ability to follow the gaze of another (Itakura & Tanaka, 1998; Okamoto et al., 2002; Tomasello, Hare, Lehmann, & Call, 2007) and the holistic (configural) processing of faces (Parr, Dove, & Hopkins, 1998; Tomonaga, 1999, 2007b); however, the evidence for holistic face processing in monkeys remains controversial (Adachi, Chou, & Hampton, in press; Dahl, Wallraven, Bulthoff, & Logothetis, 2009; Parr & Heintz, 2008; Tomonaga, 1994).

On the other hand, notable differences between humans and chimpanzees have also been identified. Face-to-face communications in chimpanzees tend to occur in brief spurts and during arousing situations, and the most typical close-range long-bout affiliative communications are tactile (e.g., grooming: Goodall, 1968; van Hooff, 1973). Humans have fine motor control of the muscles around the eyes, enabling the formation of subtle expressions in these regions (Ekman & Friesen, 1978). In contrast, these regions are hardly visible in chimpanzees because of their high eyebrow ridges and dark eyes (Kobayashi & Kohshima, 1997, 2001; Parr et al., 2007). Instead, chimpanzees have fine motor control of their lips, enabling the formation of more variable expressions in the mouth region (Chevalier-Skolnikoff, 1973; Parr et al., 2007; Vick et al., 2007). Kobayashi and Kohshima (2001) found that, compared with other primates, humans have exceptionally large white sclera that clearly contrast with the colours of their iris and skin. These authors hypothesized that human eyes have adapted to enhance gaze signals, such as eye directions. Experimental studies have shown that chimpanzees were less sensitive than human infants to the eye directions of a human experimenter when the experimenter's head was immobile (Tomasello et al., 2007), possibly because chimpanzees inspected the experimenter's face more briefly during their interactions (Carpenter & Tomasello, 1995). Additionally, in an experiment using a computerized task, chimpanzees did not show overt evidence of reflexive shifts of attention in the same direction as the eyes in a human stimulus face (Tomonaga, 2007a).

These results consistently suggested that eye regions (upper faces) are especially important in humans, compared with chimpanzees, in facial communications. Fine muscles and eye movements are accentuated by hair and colour-contrast in the eye regions in humans, signalling a variety of communicative intents. These morphologies are visually salient and thus might independently attract the attention of other individuals. However, given that chimpanzees were less sensitive than were humans to the eye direction of humans, humans might also have characteristic behavioural/attentional sensitivities to the eye regions. That is, humans might have stronger behavioural tendencies than chimpanzees to actively collect information from the eye regions.

Carpenter et al. (1995) found that, in interactive situations (facilitative of joint attention), human infants looked at the experimenter's face an average of twice as long as chimpanzees. Interestingly, in this related study, the human-raised and enculturated chimpanzees did not differ from those reared by biological mothers in this regard. Kano and Tomonaga (2009) used the eye-tracking method to measure how humans (adults) and chimpanzees viewed whole-body images of conspecifics and non-conspecifics. Both species showed highly similar patterns of scanning and both paid more attention to the face than to the other body regions depicted in the images. Additionally, both species showed, on average, an equal number of fixations on the faces. The critical difference was that the average duration of fixation on the faces was shorter in chimpanzees (300 ms) than in humans (680 ms). These differences were consistent across conspecific and non-conspecific images. However, the patterns of face scanning for each facial feature (e.g., eyes, mouth) have remained unclear because eye gaze was disproportionately directed to faces versus to the rest of the body in these prior studies.

This study thus aimed to determine the patterns of face scanning in chimpanzees. We examined the spatial (where) and temporal (when) characteristics of fixation sequences, and compared them with those of humans. We presented close-up photographs of faces to enhance our ability to observe eye movements as participants scanned the images and measured frequencies, durations and probabilities of fixations. Although widely employed in humans (Henderson, Williams, & Falk, 2005; Pelphrey et al., 2002; Walker-Smith, Gale, & Findlay, 1977; Yarbus, 1967) and monkeys (Gothard, Erickson, & Amaral, 2004; Guo, Robertson, Mahmoodi, Tadmor, & Young, 2003; Keating & Keating, 1982; Mendelson, Haith, & Goldmanrakic, 1982; Nahm, Perret, Amaral, & Albright, 1997; Sato & Nakamura, 2001), this methodology has not previously been used to investigate face scanning in apes. Humans and monkeys are known to intensely scan the main features of faces (i.e., eyes, nose, and mouth). It is also known that these species look at the eye regions longer than at other facial features and are more likely to initially inspect the eye regions. The comparisons between humans and monkeys have thus suggested qualitative similarities in face scanning. However, these previous studies were not particularly designed to directly compare human and nonhuman primates, further studies are necessary to reveal both qualitative and quantitative similarities and differences between the species using the phylogenetically closer species to humans and the comparable experimental procedures. This study thus directly compared humans and their closest evolutionary relatives, chimpanzees, using a non-invasive eye-tracking method under unrestrained conditions. The fixation sequences of humans were characterized by sequential fixations over the eyes and mouth, which appeared to somewhat resemble inverted triangular traces (Walker-Smith et al., 1977; Yarbus, 1967). The precise comparisons between species enabled us to analyze the characteristic patterns underpinning the sequential/spatiotemporal aspects of eye movements in each species.

We conducted two experiments. Experiment 1 presented both conspecific and non-conspecific faces and examined general similarities and differences in face scanning between chimpanzees and humans. We addressed three questions in Experiment 1. The first question addressed whether the patterns of scanning in chimpanzees and/or humans depended on specific responses to face stimuli; that is, we examined whether these patterns depended solely on general responses to relatively broad classes of stimuli or solely on the low-level guidance of eye movements evoked by the visual salience of image features (e.g., white sclera in humans, protruding nose in humans, protruding mouth in chimpanzees). We addressed these issues by examining the patterns of scanning, especially with respect to the main facial features (i.e., eyes nose, mouth) because previous studies on monkeys and humans have suggested that the patterns of scanning specific to face stimuli were characterized by systematic responses to these main features. We also examined consistencies in the viewing patterns for faces found by the current and previous (Kano & Tomonaga, 2009) studies involving the presentation of close-up shots of faces and full shots of whole bodies to examine the effect of scale on eye movements. The second question addressed the characteristic patterns of face scanning in chimpanzees and humans. Given the aforementioned studies, we expected that chimpanzees and humans would especially differ with regard to viewing patterns for the eye regions. The third question addressed the factors that contributed to these characteristic patterns of face scanning in chimpanzees and humans. Previous studies have suggested that these patterns might be influenced by exposure to certain types of faces (e.g., own/other race: Michel, Rossion, Han, Chung, & Caldara, 2006; reared by own/other species: Martin-Malivel & Okada, 2007) and by phylogenetic relatedness (e.g., conspecific/non-conspecific: Pascalis & Bachevalier, 1998). Indeed, characteristic patterns might reflect more general responses to limited sets of stimulus cues, such as face-like configurations. Experiment 2 presented species-specific facial expressions to chimpanzees and humans. Face-to-face interactions typically involve various emotional gestures in both species (Argyle & Cook, 1976; van Hooff, 1967). Although several direct comparisons of gaze perception in humans and chimpanzees have been conducted, as mentioned earlier, such attempts have not addressed the perception of facial expressions. Experiment 2 further examined the questions addressed in Experiment 1. More specifically, Experiment 2 addressed changes in the characteristic patterns of face scanning practiced by each species as a function of the type of expression presented. Because the quality of information in the eye and mouth regions of chimpanzees and humans differ, as mentioned earlier, we expected to find differences in their responses to the eye and mouth regions in various expressions.

Methods

We used the identical experimental framework, with the exception of the stimuli, as that used by Kano & Tomonaga (2009) (these are referred to as ‘the previous experiments” in the Methods section) to allow comparisons between the studies.

Experiment 1

Participants and apparatus. Six chimpanzees (5 females, 1 male; aged 8–31) and 18 humans (11 females, 7 males; aged 18–31; all Japanese students) participated in Experiment 1. All chimpanzees and half the humans had participated in the previous experiments. The chimpanzees were members of a social group comprised of 14 individuals living in an enriched environment with a 700-m^2 outdoor compound and an attached indoor residence (Matsuzawa, Tomonaga, & Tanaka, 2006). The outdoor compound was equipped with 15-m-high climbing frames, small streams, and various species of trees (Ochiai & Matsuzawa, 1997). Access to the outdoor compound was available to each individual every other day during the day. Daily meals included a wide variety of fresh fruits and vegetables fed throughout the day, supplemented with nutritionally-balanced biscuits (fed twice daily) and water available *ad libitum*. Both chimpanzee and human participants had extensive experience interacting with both species and were thus highly familiar with both kinds of faces. The chimpanzees, like the human participants, had extensive experience observing photographs of faces (Matsuzawa et al., 2006) and thus never responded fearfully to the faces. No food or water deprivation was practiced during the study period. Care and use of the chimpanzees adhered to the 2002 version of the Guidelines for the Care and Use of Laboratory Primates by the Primate Research Institute, Kyoto University. The experimental protocol was approved by the Animal Welfare and Care Committee of the institute. Informed consent was obtained from all human participants. Both species used the same apparatus to allow for direct comparisons. Participants sat still and unrestrained in an experimental booth and viewed a 17-inch LCD display (1280 × 1024 pixels) at a distance of approximately 60 cm. A table-mounted eye-tracker measured their eye movements using infrared corneal reflection techniques (60 Hz; Tobii X120, Tobii Technology AB; Appendix 1). As a result of the training conducted during the previous experiment, chimpanzees were already skilled at sitting still in front of an eye-tracker and looking at a fixation point that appeared on the screen. We conducted two-point and five-point calibrations for chimpanzees and humans, respectively. Calibrations were repeated until maximum accuracy was obtained, resulting in high spatial resolution for the eye tracking in both species. In the preliminary recording, six participants of each species were asked to look at a fixation point and the error values—the average distance between the intended and the recorded fixations—were measured as 0.62±0.06 of a degree (mean±SEM) for the chimpanzees and 0.52±0.05 of a degree for the humans. Refer to the previous experiment for details about calibration training and procedures.

Stimuli and procedures. Stimuli consisted of 24 colour still photographs of the faces of 17 species of non-primate mammals (giraffes, rhinos, lions, etc.; hereafter, mammals), 24 photographs of chimpanzee faces, and 24 photographs of human faces (frontal views; 12 individuals, six of whom were familiar and six of whom were unfamiliar to participants; see

Fig. 1 for examples). Both species had relatively less experience with observing mammal than chimpanzee and human faces. Half of these facial images were extracted from the whole-body images used in the previous experiments. The photographs were converted into 1000 × 800 pixels with surrounding gray frames (1280 × 1024 pixels in total). During testing, each trial was initiated by the participants looking at a fixation point that appeared at a random position on the screen. The photographs were then presented for 2 s, and participants were allowed to move their eyes freely to view the photograph. A total of 72 photographs were presented to the humans within a single day, whereas the sessions were divided among 10 days for the chimpanzees to maintain their spontaneous motivation for viewing photographs. The order in which photographs were presented was randomized within the entire session for each participant. The human participants received 500 yen after the session, and the chimpanzees obtained a small piece of apple after each trial, regardless of their viewing behaviours.

Experiment 2

Five chimpanzees (4 females, 1 male) and 9 humans (7 females, 2 males) participated in Experiment 2. One chimpanzee (adult female) was eliminated from testing because of her lack of attention to the stimuli. The stimuli consisted of 12 colour still photographs portraying species-specific chimpanzee facial expressions (three each portraying neutral faces, hoot faces, scream faces, and compressed-lip faces; see Parr et al., 2007 for descriptions of these expressions) and 12 monochrome still photographs of a standardized set of human facial expressions (three each portraying neutral faces, happy faces, fearful faces, and angry faces; taken from Ekman & Friesen, 1978; see Fig. 3 for examples). Chimpanzee facial expressions were obtained by videotaping social interactions at another chimpanzee colony and then isolating the frames containing the expressions reflecting peak intensity. The photographs of chimpanzee and human expressions were converted into 1000 × 800 pixels and 600 × 800 pixels, respectively (no background in human photographs), with the surrounding gray frames (1280 × 1024 pixels in total). Facial expressions were presented for 2 s to chimpanzees. Given the slower inspection of each facial feature by humans (see below) than by chimpanzees, which was found in Experiment 1, faces were presented for 5 s to humans in order to leave sufficient time for them to explore each facial feature. Each trial presented a photograph that was randomly drawn from each type of facial expression. The remaining procedures were the same as those used in Experiment 1.

Data Analysis

Trials in which participants did not view the monitor for more than 300 ms were eliminated from the analysis, resulting in a loss of 4.1% of the data obtained from chimpanzees (no data were lost for humans). During Experiment 2, we repeated these trials (8.3% of all the data obtained from chimpanzees) after the end of session, resulting in no loss of data for either species. To define areas of interest (AOI), the scenes were divided into face regions and the rest of the scenes. The face regions were further divided into the eye, nose (referred as the "mid-face" for mammal faces), mouth, and other regions (see Fig. 1 for examples). To avoid errors in gaze estimations, AOIs were drawn slightly larger than the actual outlines (approximately 20 pixels on the edges). A fixation was scored if the gaze remained stationary (within a radius of 50 pixels) for at least 75 ms (more than 5 measurement samples). Otherwise, the recorded sample was defined as part of a saccade. We excluded the samples recorded during the first 200 ms, thereby eliminating fixations that followed the offset of the fixation spot. We used four dependent variables in this study: cumulative viewing time, number (frequency) of fixations, average fixation duration, and the probability of fixations (as a function of fixation order). The probability of fixation was calculated as the proportion of photographs in which a certain AOI was the target at a particular point in the fixation order. In Experiment 2, the number of fixations was calculated as a proportion of the total number of face fixations to correct for the differences between the two subject species in terms of presentation duration. For the statistical analyses, we distinguished within-species from between-species comparisons. For within-species comparisons, we tested for differences in the viewing patterns for each AOI within each subject species. For between-species comparisons, we tested for the interactions between subject species and AOIs in viewing patterns. In ANOVAs, subject species and other independent variables (facial features, fixation order, facial expressions, and presentation sections) served as between- and within-subject factors, respectively. In cases in which the assumption of homogeneity of variance was violated, the Greenhouse-Geisser correction was applied, and corrected *p* values were calculated. We used *post-hoc t*-tests tests for within- and between-species comparisons, as well as Bonferroni's corrections with the *alpha* level set at 0.05 for the number of comparisons. We conducted all analyses independently for each stimulus species (i.e., chimpanzees, humans, and mammals), and thus did not include stimulus species as a factor in the ANOVA because facial morphologies (the proportion of each face occupied by each feature) differed somewhat by stimulus species. In Experiment 2, we normalized the data for each expressive face according to the proportions characterizing the neutral faces of the same stimulus species to enable comparisons among the viewing patterns associated with different facial expressions. We then independently compared the data obtained in response to each stimulus species according to facial expression.

Results

Experiment 1

We obtained several results consistent with those of previous experiments (Kano & Tomonaga, 2009) involving the presentation of the whole-body images of the same animals. First, both species fixated on the face region more frequently than on the rest of the scenes for all species (chimpanzee faces: 4.5 vs. 1.3 times, 4.2 vs. 0.14 times; human faces: 3.8 vs. 2.0 times, 3.7 vs. 0.23 times; mammal faces: 4.6 vs. 1.5 times, 4.1 vs. 0.23 times, for chimpanzee and human participants respectively; $p < 0.05$). Second, chimpanzees and humans did not differ significantly in the numbers of fixations on the face regions of any species ($p > 0.05$). Third, the average fixation duration for faces was shorter in chimpanzees than in humans (see Table 1).

The new findings emerging from this study involved similarities and differences between the species with regard to the scanning patterns for each facial feature. We focussed on the number of fixations and on the probability of fixations as a function of fixation order (i.e., eye movement path; the variables represented by the yellow lines in Fig. 1) in deriving the following results because we obtained similar results when the other two variables were included in the analysis (i.e., cumulative viewing time and average fixation duration; the variables indicated by the orange circles in Fig. 1; see Table 1 for these results). Chimpanzees and humans fixated on the main facial features (i.e., eyes, nose, and mouth) more frequently than they fixated on the other regions of chimpanzee faces (Fig. 2a; $t_5 = 5.27, p = 0.003$, $t_{17} = 24.40, p < 0.001$, respectively) and of human faces ($t_5 = 6.70, p = 0.001$, $t_{17} = 17.71, p < 0.001$, respectively). Although this pattern of results did not emerge in chimpanzees with regard to mammal faces ($t_5 = 1.51, p = 0.19$) but did emerge in humans with regard to mammal faces ($t_{17} = 8.77, p < 0.001$), this phenomenon probably derived from the frequent inspection of the other regions (e.g. the mane of a lion, the horn of a rhino) by both chimpanzees and humans. The ANOVA focussing on the main facial features revealed a significant interaction between subject species and facial features (chimpanzee faces: $F_{1.1, 24} = 11.5, p < 0.001, \eta_p^2 = 0.34$; human faces: $F_{1.2, 26} = 6.03, p = 0.005, \eta_p^2 = 0.21$; mammal faces: $F_{1, 22} = 44.7, p < 0.001, \eta_p^2 = 0.67$). *Post-hoc* comparisons between the species revealed that humans fixated on the eye region more frequently than did chimpanzees (Fig. 2a; chimpanzee faces: $t_{22} = 2.93$, $p = 0.008$; human faces: $t_{22} = 2.73$, $p = 0.012$; mammal faces: $t_{22} = 5.30$, $p < 0.001$), and chimpanzees fixated on the mouth region more frequently than did humans (chimpanzee faces: $t_{22} = 4.21$, $p < 0.001$; human faces: $t_{22} = 3.34$, $p = 0.003$; mammal faces: $t_{22} = 3.62$, $p = 0.002$). *Post-hoc* comparisons within species revealed that, unlike the humans (chimpanzee faces: $t_{17} = 9.29$, $p < 0.001$; human faces: $t_{17} = 7.79$, $p < 0.001$; mammal faces: $t_{17} = 13.9$, $p < 0.001$), the chimpanzees did not fixate on the eye regions more frequently than they fixated on the mouth regions (chimpanzee faces: $t_{17} = 1.44$, $p = 0.20$; human faces: $t_{17} = 3.92$, $p = 0.011$; mammal faces: $t_{17} = 1.26$, $p = 0.26$).

Figure 2b illustrates the temporal character of the aforementioned differences between the species for the eyes and mouth, respectively. An ANOVA revealed significant interactions between subject species and fixation order for the eyes (chimpanzee faces: $F_{3, 66} = 3.63$, $p = 0.017$; human faces: $F_{2.0, 45} = 4.36$, $p = 0.007$, $\eta_p^2 = 0.16$; mammal faces: $F_{3, 66} = 11.68$, $p = 0.001$, $\eta_p^2 = 0.34$) and for the mouth (chimpanzee faces: $F_{3, 66} = 3.82$, $p = 0.014$, $\eta_p^2 = 0.14$; human faces: $F_{3, 66} = 4.26$, $p = 0.008$, $\eta_p^2 = 0.16$; mammal faces: $F_{3, 66} = 9.13$, $p < 0.001$, $\eta_p^2 = 0.29$). We then conducted *post-hoc* tests for each fixation order. Both species scanned the eyes, followed by the mouth, as indicated by the higher probability of first fixations on the eye region than on the mouth region ($p < 0.05$ for both species viewing faces of all species). Additionally, the probabilities for first fixations did not differ significantly between the species for the eyes or mouth ($p > 0.05$). However, chimpanzees were less likely than were humans to fixate on the eye region during later fixations ($p < 0.01$). Chimpanzees were significantly more likely than humans to fixate on the mouth region as their second fixation ($p < 0.01$), but not as their third or fourth fixation ($p > 0.05$). That is, although chimpanzees and humans both began their fixation sequences with the eye regions, humans were more likely than chimpanzees to subsequently re-fixate on the eye regions, whereas the chimpanzees were more likely than humans to subsequently shift their gazes to the mouth region.

We should rule out two possible effects of the presentation procedures on the results. First, the chimpanzees might have habituated to the faces more quickly than did the humans, enabling them to scan the faces more rapidly. We divided the entire sessions into three sections and compared the number of fixations on the face regions among these three sections. However, an ANOVA did not find any significant interaction between subject species and section (chimpanzee faces: $F_{2,44} = 1.62$, $p = 0.20$, $\eta_p^2 = 0.069$; human faces: $F_{2,44} = 0.376$, $p = 0.19$, $\eta_p^2 = 0.072$; mammal faces: $F_{2,44} = 1.96$, $p = 0.15$, $\eta_p^2 = 0.082$). Second, the presentation duration in this study (2 s) might have been too short for the humans, leaving insufficient time to explore facial features other than eyes. Thus, the same five human participants viewed half of the same human and chimpanzee faces again, but with a presentation duration of 10 s. We found that the moderate decrease/increase in the probability of fixation on the eye/mouth regions by humans (as seen in Fig. 2b) remained during this longer presentation and that the decrease/increase stabilized at the seventh fixation at a probability of 0.4-0.6 for the eye region and of 0.0-0.2 for the mouth region. Thus, we confirmed that the aforementioned differences between species were not related to presentation order or presentation duration.

Experiment 2

Figure 3 illustrates the attentional responses of chimpanzees and humans to species-specific facial expressions of both species. Only the eye and mouth regions of the chimpanzee and human faces (regions that change in facial expressions) were included in this analysis to avoid redundancy. We first conducted ANOVAs (subject species × facial features × facial expressions) and found a significant interaction among the three factors (chimpanzee faces: $F_{3, 36} = 4.14$, $p = 0.013$, $\eta_p^2 = 0.25$; human faces: $F_{3, 36} = 1.62$, $p = 0.007$, $\eta_p^2 = 0.28$). We then conducted ANOVAs (subject species × facial expressions) for eye and mouth regions and found that humans viewed the eye regions more frequently than did chimpanzees (chimpanzee faces: $F_{1, 12} = 8.41$, $p < 0.001$, $\eta_p^2 = 0.89$; human faces: $F_{1, 12} = 15.2$, $p < 0.001$, $\eta_p^2 = 0.94$), and chimpanzees viewed the mouth regions more frequently than did humans (chimpanzee faces: $F_{1, 12} = 40.1$, $p < 0.001$, $\eta_p^2 = 0.96$; human faces: $F_{1, 12} = 17.1$, $p < 0.001$, $\eta_p^2 = 0.92$). Although ANOVAs (facial features × facial expressions) revealed that chimpanzees differentiated among the facial expressions in terms of the proportion of fixations to eye or mouth regions (chimpanzee faces: $F_{3, 12} = 8.37$, $p = 0.003$, $\eta_p^2 = 0.67$; human faces: $F_{3, 12} = 11.0$, $p = 0.001$, $\eta_p^2 = 0.73$), this was not the case for humans (chimpanzee faces: $F_{1.3, 11} = 0.61$, $p = 0.615$, $\eta_p^2 = 0.071$; human faces: $F_{3, 24} = 0.29$, $p = 0.82$, $\eta_p^2 = 0.036$). To examine the effect of the different presentation durations used for humans (5 s) and chimpanzees (2 s), we conducted the same analyses for the first 2 s of the presentations to humans. However, we confirmed a similar tendency in humans (intense and persistent eye-viewing) in this analysis. These results indicate that although chimpanzees changed their scanning patterns in response to facial actions, humans maintained their intense focus on the eye regions across expressions. This difference between the species was consistent across chimpanzee and human faces.

Discussion

To our knowledge, this is the first reported study to obtain comparative data on face scanning in chimpanzees and humans, thereby offering methodological advances for examining the evolution of face perception and facial communication. The observed scanning patterns were not dependent solely on the visually salient features in the faces, as evidenced by the systematic similarities and differences between the species in their scanning patterns for the main features of faces. We confirmed the robust consistency between the current and previous study (Kano & Tomonaga, 2009) involving the presentation of close-up shots of faces and full shots of whole bodies. These results indicated that chimpanzees and humans exhibited patterns of scanning specific to face stimuli. Several qualitative similarities highlighted the homologous nature of face scanning of chimpanzees and humans: both demonstrated intense scanning of the main facial features (except in mammal faces; see Results) and the same order of inspection for each facial feature (from the eyes to mouth). Importantly, notable quantitative differences between chimpanzees and humans also emerged: the prolonged eye-viewing by humans (the sequential re-fixations on the eyes), the quick, vertical scanning of faces by chimpanzees (immediate shifts of eye gaze from the eyes to mouth). If we emphasized the differences in the inverted triangular sequences of fixations over the eyes and mouth exhibited by humans, the sequences of fixations exhibited by chimpanzees would appear to resemble vertical segments of lines (as seen in Fig. 1).

These similarities and differences between the species were consistent across conspecific faces, non-conspecific faces, and even phylogenetically distant mammal faces. Recall that the chimpanzees and humans in this study were highly familiar with chimpanzee and human faces, but relatively unfamiliar with the mammal faces (see Methods). Thus, these patterns did not appear to reflect exposure to a certain type of faces or to derive from phylogenetic relatedness, but rather seemed to involve more general responses to the face-like configurations. Consistent with Experiment 1, Experiment 2 showed that chimpanzees and humans viewed the mouth and eye regions, respectively, more frequently than did the other species. Experiment 2 also demonstrated that although chimpanzees changed their scanning patterns in responses to changes in facial expressions, and humans maintained intense eye-viewing across the expressions, these differences were consistent across chimpanzee and human faces, which also matches the results of Experiment 1.

In general, the patterns of face scanning in humans were characterized by prolonged eye-viewing irrespective of facial expression, whereas those in chimpanzees were characterized by quick, vertical scanning of faces, frequent inspection of the mouth regions, and responsiveness to facial actions. The characteristics of human scanning patterns indicate active viewing of the eye regions that is independent of facial morphology and expression. On the other hand, although the overall patterns of eye movements in chimpanzees were characterized by specific responses to face stimuli (see above), some of the characteristics might be attributable to either general responses to relatively broad classes of stimuli or to the low-level guidance of eye movements by the visual salience of image features. For example, the rapid scanning in chimpanzees might reflect their general ability to scan scenes. In the previous experiment (Kano & Tomonaga, 2009), chimpanzees exhibited a more rapid shift in fixation location than did humans in response to overall scenes (i.e., not only for faces but also for the rest of the scene). Additionally, the frequent inspection of mouth regions demonstrated by chimpanzees might represent passive viewing of the visually salient regions, reflecting their reluctance to view eyes, rather than active viewing of the informative regions. The mouth regions are not only informative with regard to emotions but also visually salient, especially in chimpanzee faces. Thus, although both species seem highly motivated to view faces (chimpanzees to a lesser extent), the attention of chimpanzees might be less focused and more spatially dispersed.

The unique eye morphologies in humans appear to be adapted to enhance particular signals such as gaze direction (Kobayashi & Kohshima, 2001). Thus, the active viewing of eyes by humans might be a behavioural adaptation to enhance unique forms of facial communications that use the eyes. Likewise, the characteristic patterns of face scanning in chimpanzees might also have been adapted for communicative purposes, especially for the purpose of the rapid and efficient retrieval of emotional information from faces. However, we doubt the latter possibility because the characteristic in chimpanzee scanning patterns might not derive from the active viewing of particular facial features, as mentioned earlier. Furthermore, the salient (large) mouth morphology in chimpanzees appears to be adapted for purposes other than communicative signalling, such as food processing (Lambert, 1999).

Irrespective of the selective pressures that have shaped the characteristic patterns of face scanning in chimpanzees and humans, it should be noted that each pattern seems to offer advantages for species-specific forms of facial communications. The persistent and prolonged eye-viewing exhibited by humans might enable the constant retrieval of subtle information conveyed by the eye region. As mentioned earlier, facial muscles and unique eye morphologies contribute to the formation of various subtle expressions in the eye regions of humans. Thus, it would appear to be advantageous for humans to remain focused on the eye region to recognize these species-specific forms of expression in the eye region. Indeed, it is known that autistic children, who experience difficulties in forming normal social and emotional relationships with people, pay less attention to the eye regions than do children with typical developmental courses (Dalton et al., 2005; Dawson et al., 2004; Klin, Jones, Schultz, Volkmar, & Cohen, 2002; Pelphrey et al., 2002). The quick scanning of faces by chimpanzees, on the other hand, might enable them to coarsely but quickly retrieve overall information (e.g., identity, emotion) from faces. As mentioned earlier, compared to humans, face-to-face communications among chimpanzees tend to occur in brief spurts and during arousing situations. In addition, the mouth regions constitute the main sources of emotional expression in chimpanzees. Thus, the characteristic patterns of face scanning in chimpanzees—the quick, vertical scanning of faces, the frequent inspection of the mouth, and the responsiveness to facial actions—would appear to be advantageous for efficiently retrieving general as well as emotional information from faces. The rapid inspection of faces might also be advantageous for chimpanzees with regard to appeasing the other individuals in that it is more likely that prolonged eye contact functions as a threat signal in non-human primates than in humans (Gomez, 1996; Thomsen, 1974). These findings suggest an intimate connection between face scanning and species-specific forms of facial communications in chimpanzees and humans, rendering their respective eye movement strategies functional in interactions with conspecifics.

The direct causes for these characteristic scanning patterns in chimpanzees and humans remain unclear; further studies are necessary. For example, the reluctance of chimpanzees to view eyes might be due to the less powerful incentive for chimpanzees, compared with humans, to collect eye information, but might also be due to the more powerful incentive for chimpanzees to avoid eye contact. The active viewing of eyes by humans might have developed to enhance species-specific forms of emotional communication in which the eye regions play unique roles, as discussed above. It is also possible that such tendencies in humans developed in the service of non-emotional instrumental communication in which facial expressions, including the eyes, serve language-like functions. Eye movements signal not only emotion but also direction of attention or interest. The importance of joint attentional and communicative interactions in humans might have shaped the unique morphologies of the eyes (Tomasello et al., 2007) and also the behavioural tendency to actively collect eye information. From the ontogenetic perspective, the active viewing of eyes by humans might derive from biologically determined patterns reflecting certain selective pressures in evolution, but might also constitute patterns learned during the course of development. Studies in human infants have shown that infants dramatically increased fixations on the eye region at about seven weeks of age (Haith, Bergman, & Moore, 1977). It is known that the differential patterns involving direct/averted gaze emerge early in life (within a few days/weeks) in humans (Farroni, Csibra, Simion, & Johnson, 2002), chimpanzees (Myowa-Yamakoshi, Tomonaga, Tanaka, & Matsuzawa, 2003), and monkeys (Mendelson et al., 1982). However, precise measurements of eye movements have not yet been conducted in infant chimpanzees, and direct comparisons involving these primate infants are necessary to clarify the evolution and development of eye contact.

Because the experiments in this study were conducted in the absence of specific communicative contexts, these results most likely reflect general species differences that are relatively consistent across various communicative situations. How these characteristic patterns of face scanning in chimpanzees and humans respond to various communicative contexts remains unknown, and further studies are necessary. Unlike monkey species and similar to humans, chimpanzees are known to frequently engage in relatively long bouts of eye contact in affiliative interactions (Goodall, 1986). Chimpanzees tend to alternate gazing at food and at the experimenter to obtain food, perhaps with communicative intent (Leavens & Hopkins, 1998).

The absence of communicative contexts in these experiments might have promoted the demonstration of general responses to the faces rather than specific responses to particular faces (e.g., familiar/unfamiliar species, phylogenetically close/distant species), as well as possibly enhancing differences between species rather than differences within each species. Although the inspection of the data obtained from the individual participants in this study suggested that each individual manifested several behavioural trends that were consistent across experiments, including those conducted previously (Kano & Tomonaga, 2009), this analysis did not suggest specific intraspecies differences in terms of social rank, age, sex, etc. Likewise, previous studies in humans have revealed that East Asians tended to look at the eye regions for shorter durations than did Caucasians when presented with photographs of East Asian and Caucasian faces (Blais, Jack, Scheepers, Fiset, & Caldara, 2008). Given that the human participants in this study (all Japanese) showed more attention to the eye regions than did the chimpanzee participants, the aforementioned cultural difference appears to be less pronounced than the interspecies difference between chimpanzees and humans.

Several similarities between humans and chimpanzees, such as the intense scanning of main facial features and the order in which each facial feature was inspected, are also consistent with the studies on monkeys (Ghazanfar, Nielsen, & Logothetis, 2006; Gothard et al., 2004; Guo et al., 2003). Of special interest in this regard is the characteristic pattern of human face scanning involving sequential re-fixations on the eye region. This pattern is consistent with results of previous human studies (Althoff & Cohen, 1999; Henderson et al., 2005; Walker-Smith et al., 1977). Humans and monkeys are known to look at the eye region longer than at the mouth region, as mentioned earlier. However, this was not the case for the chimpanzees in this study, probably due to the frequent fixations on the mouth region by the chimpanzees. Interestingly, preliminary comparisons with previously published data on monkeys suggested that humans viewed the eye region for the longest durations among these three primate species, that chimpanzees viewed the eye region for as long as did monkeys, and that chimpanzees viewed the mouth region for the longest durations (compare Table 1 with, for example, Guo et al., 2003). Perhaps the important aspects of human face scanning include the prolonged eye-viewing (or the sequential re-fixations on the eye regions) and not simply the dominance of the eyes over the mouth as measured in total viewing time. Additional studies are necessary to clarify this issue.

Acknowledgements

This research was financially supported by the Japan Society for the Promotion of Science (JSPS) and the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan Grants-in-Aid for Scientific Research (nos. 16002001, 19300091, 20002001, 212299) and the JSPS/MEXT global COE programs (D07 and A06). We thank Drs T. Matsuzawa, M. Tanaka, I. Adachi, S. Hirata, and Y. Hattori for their help and invaluable comments. We also thank the Centre for Human Evolution Modelling Research at the Primate Research Institute for the daily care of the chimpanzees.

References

Adachi, I., Chou, D. P., & Hampton, R. R. (in press). Thatcher effect in monkeys demonstrates conservation of face perception across primates. *Current Biology*.

Althoff, R. R., & Cohen, N. J. (1999). Eye-movement-based memory effect: A reprocessing effect in face perception. *Journal of Experimental Psychology: Learning Memory and Cognition, 25*(4), 997-1010.

Argyle, M., & Cook, M. (1976). *Gaze and mutual gaze*. Cambridge: Cambridge University Press

Bard, K. A., Myowa-Yamakoshi, M., Tomonaga, M., Tanaka, M., Costall, A., & Matsuzawa, T. (2005). Group differences in the mutual gaze of chimpanzees (*Pan troglodytes*). *Developmental Psychology, 41*(4), 9.

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS ONE, 3*(8).

Carpenter, M., & Tomasello, M. (1995). Joint attention and imitative learning in children, chimpanzees, and enculturated chimpanzees. *Social Development, 4*(3), 217-237.

Chevalier-Skolnikoff, S. (1973). Facial expression of emotion in nonhuman primates. In P. Ekman (Ed.), *Darwin and facial expression: A century of research in review* (pp. 11-89). New York: Academic Press.

Dahl, C. D., Wallraven, C., Bulthoff, H. H., & Logothetis, N. K. (2009). Humans and macaques employ similar face-processing strategies. *Current Biology, 19*(6), 509-513.

Dalton, K. M., Nacewicz, B. M., Johnstone, T., Schaefer, H. S., Gernsbacher, M. A., Goldsmith, H. H., et al. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nature Neuroscience, 8*, 519-526.

Darwin, C. (1999). *The expression of the emotions in man and animals (P. Ekman, Ed)*. London: Fontana. (Original work published 1872).

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., et al. (2004). Early social attention impairments in autism: Social orienting, joint attention, and attention to distress. *Developmental Psychology, 40*(2), 271-282.

Ekman, P., & Friesen, W. V. (1978). *The facial action coding system (FACS): A technique for the measurement of facial action*. Palo Alto: CA: Consulting Psychologists Press.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences, 99*(14), 9602.

Ghazanfar, A. A., Nielsen, K., & Logothetis, N. K. (2006). Eye movements of monkey observers viewing vocalizing conspecifics. *Cognition, 101*(3), 515-529.

Gomez, J. C. (1996). Ostensive behavior in great apes: The role of eye contact. In *Reaching into thought: The minds of the great apes* (pp. 131-151).

Goodall, J. (1968). The behaviour of free-living chimpanzees in the Gombe stream reserve. *Animal Behaviour Monographs, 1*(3), 161-311.

Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behavior*. Cambridge, MA: Belknap Press of Harvard University Press.

Gothard, K. M., Erickson, C. A., & Amaral, D. G. (2004). How do rhesus monkeys (*Macaca mulatta*) scan faces in a visual paired comparison task? *Animal Cognition, 7*(1), 25-36.

Guo, K., Robertson, R. G., Mahmoodi, S., Tadmor, Y., & Young, M. P. (2003). How do monkeys view faces?—a study of eye movements. *Experimental Brain Research, 150*(3), 363-374.

Haith, M. M., Bergman, T., & Moore, M. J. (1977). Eye contact and face scanning in early infancy. *Science, 198*(4319), 853-855.

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition, 33*(1), 98-106.

Itakura, S., & Tanaka, M. (1998). Use of experimenter-given cues during object-choice tasks by chimpanzees (*Pan troglodytes*), an orangutan (*Pongo pygmaeus*), and human infants (*Homo sapiens*). *Journal of Comparative Psychology, 112*(2), 119-126.

Kano, F., & Tomonaga, M. (2009). How chimpanzees look at pictures: a comparative eye-tracking study. *Proceedings of the Royal Society B: Biological Sciences, 276*(1664), 1949-1955.

Keating, C. F., & Keating, E. G. (1982). Visual scan patterns of rhesus monkeys viewing faces. *Perception, 11*(2), 211-219.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry, 59*(9), 809-816.

Kobayashi, H., & Kohshima, S. (1997). Unique morphology of the human eye. *Nature, 387*(6635), 767-768.

Kobayashi, H., & Kohshima, S. (2001). Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye. *Journal of Human Evolution, 40*(5), 419-435.

Lambert, J. E. (1999). Seed handling in chimpanzees (*Pan troglodytes*) and redtail monkeys (*Cercopithecus ascanius*): implications for understanding hominoid and cercopithecine fruit-processing strategies and seed dispersal. *American Journal of Physical Anthropology, 109*(3), 365-386.

Leavens, D. A., & Hopkins, W. D. (1998). Intentional communication by chimpanzees: A cross-sectional study of the use of referential gestures. *Developmental Psychology, 34*(5), 813.

Martin-Malivel, J., & Okada, K. (2007). Human and chimpanzee face recognition in chimpanzees (*Pan troglodytes*): role of exposure and impact on categorical perception. *Behavioral Neuroscience, 121*(6), 1145-1154.

Matsuzawa, T., Tomonaga, M., & Tanaka, M. (2006). *Cognitive development in chimpanzees*. Tokyo: Springer.

Mendelson, M. J., Haith, M. M., & Goldmanrakic, P. S. (1982). Face scanning and responsiveness to social cues in infant rhesus monkeys. *Developmental Psychology, 18*(2), 222-228.

Michel, C., Rossion, B., Han, J., Chung, C. S., & Caldara, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychological Science, 17*(7), 608-615.

Myowa-Yamakoshi, M., Tomonaga, M., Tanaka, M., & Matsuzawa, T. (2003). Preference for human direct gaze in infant chimpanzees (*Pan troglodytes*). *Cognition, 89*(2), 113-124.

Nahm, F. K. D., Perret, A., Amaral, D. G., & Albright, T. D. (1997). How do monkeys look at faces? *Journal of Cognitive Neuroscience, 9*(5), 611-623.

Ochiai T, Matsuzawa T (1997) Planting trees in an outdoor compound of chimpanzees for an enriched environment. *Proceedings of the Third International Conference on Environmental Enrichment*. Orlando, Florida, 355–364

Okamoto, S., Tomonaga, M., Ishii, K., Kawai, N., Tanaka, M., & Matsuzawa, T. (2002). An infant chimpanzee (*Pan troglodytes*) follows human gaze. *Animal Cognition, 5*(2), 107-114.

Parr, L. A., Dove, T., & Hopkins, W. D. (1998). Why faces may be special: evidence of the inversion effect in chimpanzees. *Journal of Cognitive Neuroscience, 10*(5), 615-622.

Parr, L. A., & Heintz, M. (2008). Discrimination of faces and houses by rhesus monkeys: the role of stimulus expertise and rotation angle. *Animal Cognition, 11*(3), 467-474.

Parr, L. A., Waller, B. M., Vick, S. J., & Bard, K. A. (2007). Classifying chimpanzee facial expressions using muscle action. *Emotion, 7*(1), 172-181.

Pascalis, O., & Bachevalier, J. (1998). Face recognition in primates: a cross-species study. *Behavioural Processes, 43*(1), 87-96.

Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., & Piven, J. (2002). Visual scanning of faces in autism. *Journal of Autism and Developmental Disorders, 32*(4), 249-261.

Sato, N., & Nakamura, K. (2001). Detection of directed gaze in rhesus monkeys (*Macaca mulatta*). *Journal of Comparative Physiology, 115*(2), 115-121.

Thomsen, C. E. (1974). Eye contact by non-human primates toward a human observer. *Animal Behaviour, 22*, 144-149.

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. *Journal of Human Evolution, 52*(3), 314-320.

Tomonaga, M. (1994). How laboratory-raised Japanese monkeys (*Macaca fuscata*) perceive rotated photographs of monkeys: Evidence for an inversion effect in face perception. *Primates, 35*(2), 155-165.

Tomonaga, M. (1999). Inversion effect in perception of human faces in a chimpanzee (*Pan troglodytes*). *Primates, 40*(3), 417-438.

Tomonaga, M. (2007a). Is chimpanzee (*Pan troglodytes*) spatial attention reflexively triggered by the gaze cue? *Journal of Comparative Psychology, 121*(2), 156-170.

Tomonaga, M. (2007b). Visual search for orientation of faces by a chimpanzee (*Pan troglodytes*): face-specific upright superiority and the role of facial configural properties. *Primates, 48*(1), 1-12.

van Hooff, J. A. R. A. M. (1967). The facial displays of the catanhine monkeys and apes. In D. Morris (Ed.), *Primate Ethology* (pp. 7-68). Piscataway, NJ: Aldine Transaction.

van Hooff, J. A. R. A. M. (1973). A structual analysis of the social behaviour of a semi-captive group of chimpanzees. In M. von Cranach & I. Vine (Eds.), *Social communication and movement* (pp. 75-162). London: Academic Press.

van Lawick-Goodall, J. (1967). Mother-offspring relationships in free-ranging chimpanzees. In D. Morris (Ed.), *Primate Ethology* (pp. 287-346). Piscataway, NJ: Aldine Transaction.

Vick, S. J., Waller, B. M., Parr, L. A., Pasqualini, M. C. S., & Bard, K. A. (2007). A cross-species comparison of facial morphology and movement in humans and chimpanzees using the facial action coding system (FACS). *Journal of Nonverbal Behavior, 31*(1), 1-20.

Walker-Smith, G. J., Gale, A. G., & Findlay, J. M. (1977). Eye movement strategies involved in face perception. *Perception, 6*(3), 313.

Yarbus, A. L. (1967). Eye movements during perception of complex objects. In *Eye Movements and Vision* (pp. 171-211). New York: Plenum Press.

Figure Legends

Figure 1. (a) Examples of face stimuli presented. The defined regions of interests are indicated by blue lines. (b) Examples of eye movements by chimpanzees and (c) by humans. Each circle represents a fixation that is linked to the adjacent fixation by a line. A longer fixation is drawn as a larger circle. Faces were presented for 2 s. See Movie 1 for a demonstration of these eye movements.

Figure 2. The similarities and differences in face-scanning patterns between chimpanzees and humans. (a) The number of fixations (mean + SEM) on each feature. (b) The probability of fixation as a function of fixation order.

Figure 3. The number of fixations as a proportion of the total numbers of face fixations (mean + SEM) on each feature of the facial expressions for chimpanzees and humans. The data were normalized to correct for the differences between facial expressions with regard to the proportions of features.

Figures

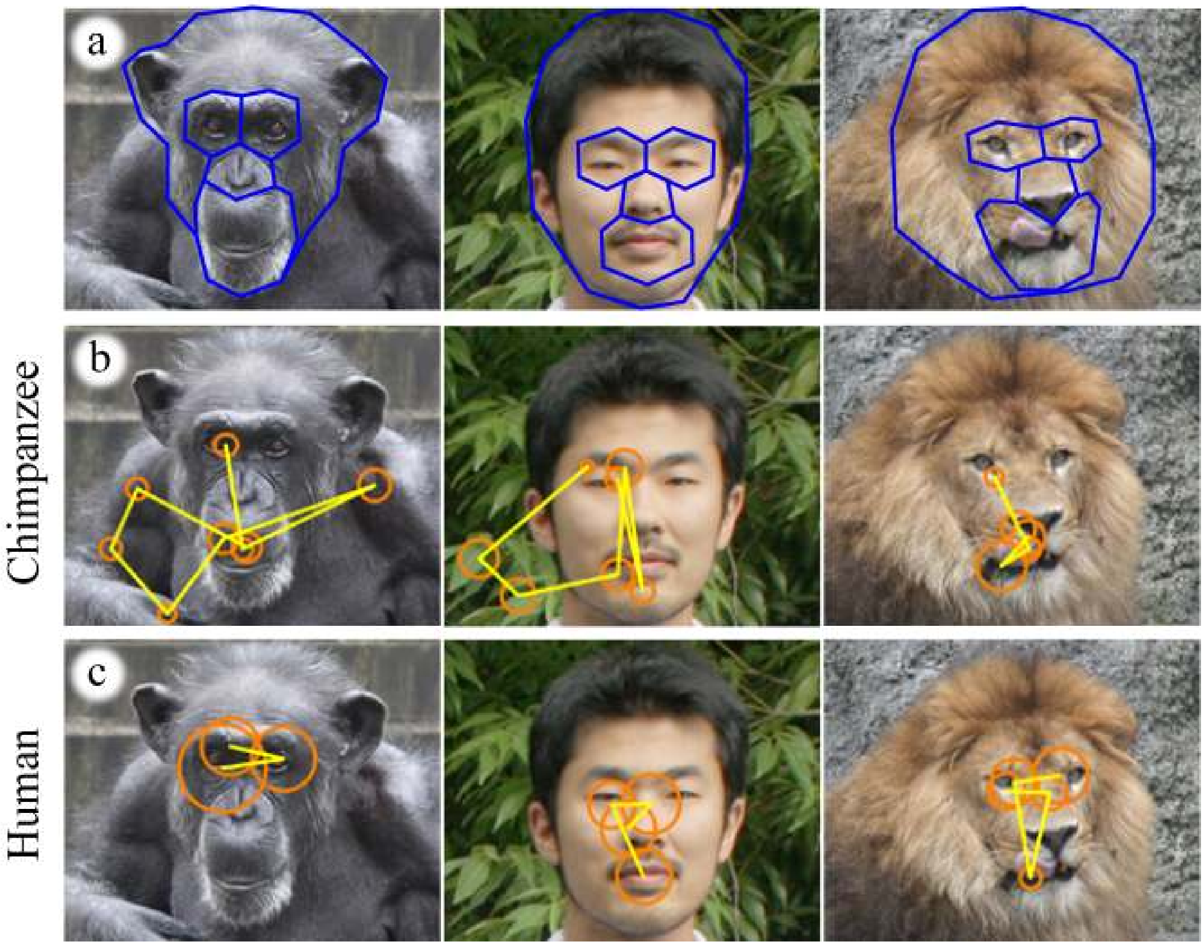


Figure 1

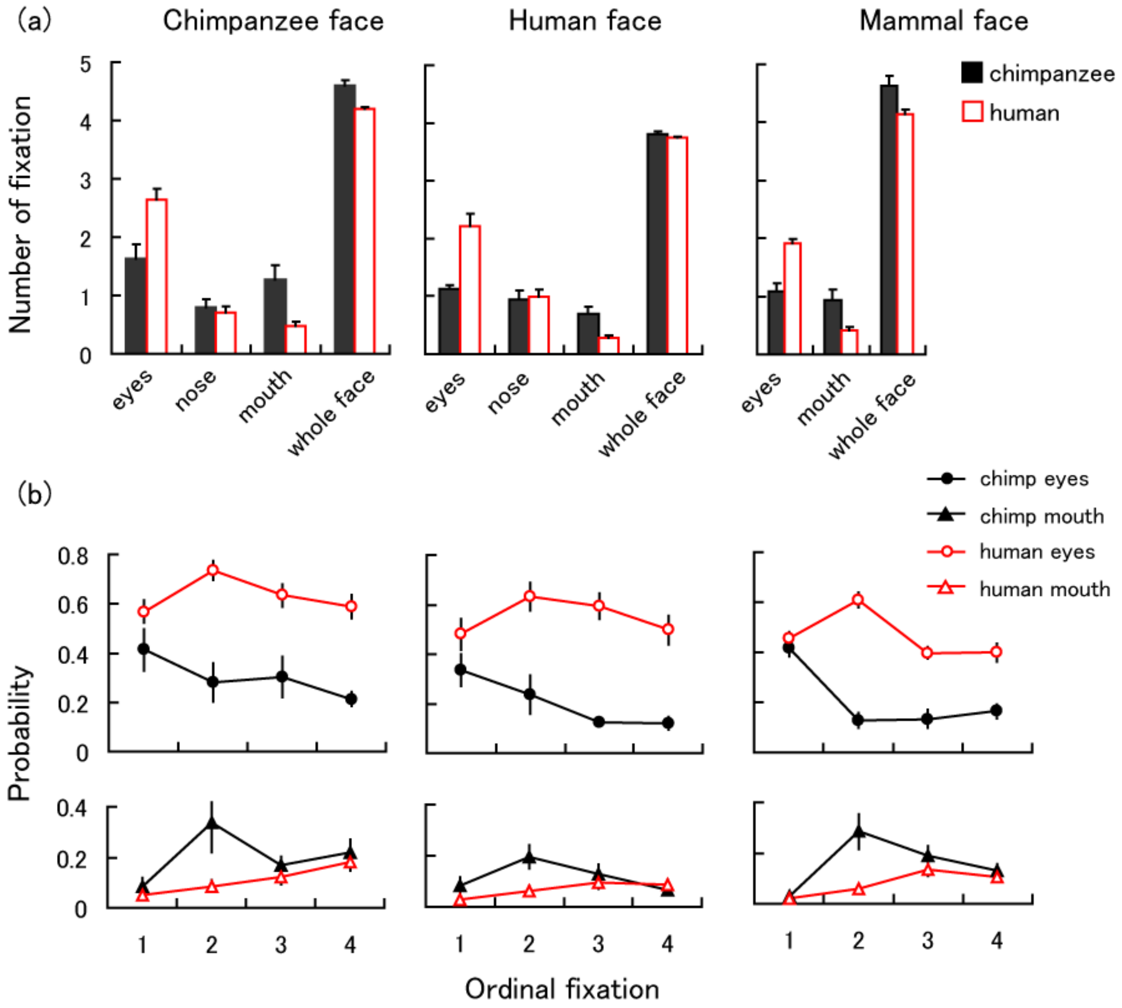
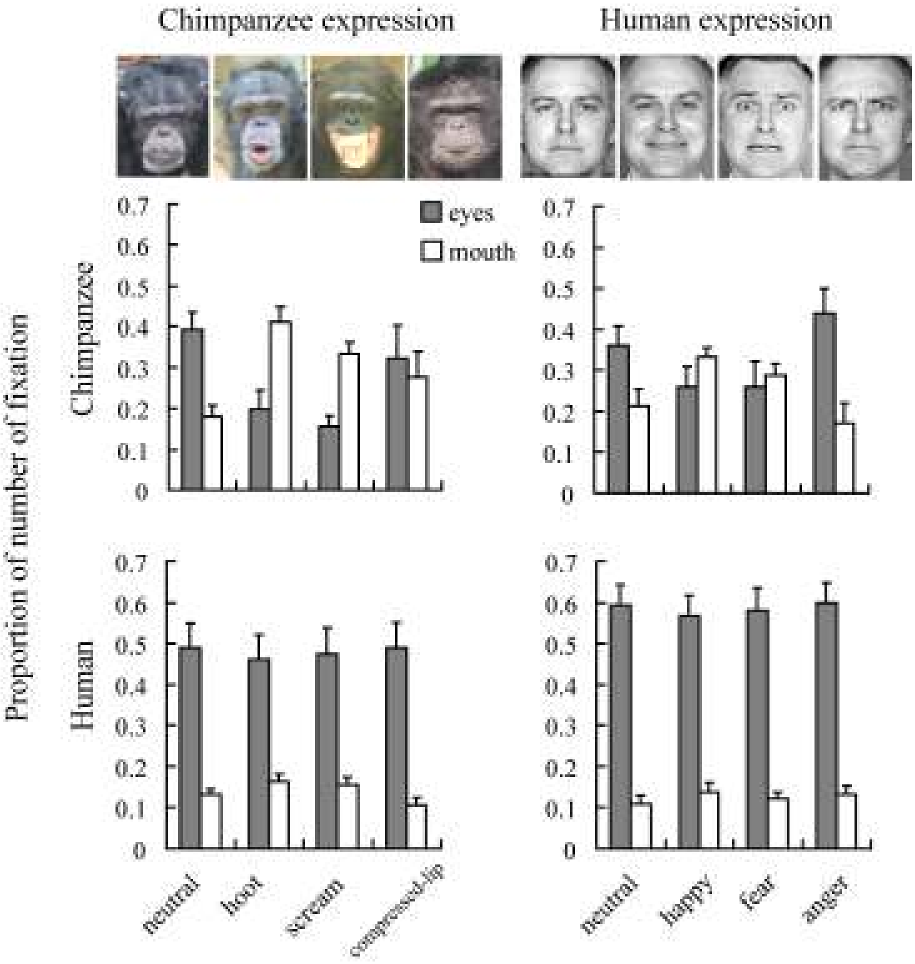

(a)
Chimpanzee face
Human face
Mammal face
chimpanzee
human
Number of fixation
5
4
3
2
1
0
eyes
nose
mouth
whole face
eyes
nose
mouth
whole face
eyes
mouth
whole face
(b)
chimp eyes
chimp mouth
human eyes
human mouth
Probability
0.8
0.6
0.4
0.2
0
0.4
0.2
0
1
2
3
4
Ordinal fixation

Figure 2



656
657 Figure 3