How do chimpanzees, *Pan troglodytes*, 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 nonconspecific faces under the same experimental conditions. Chimpanzees and humans showed 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 show sequential refixations 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 nonconspecific 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 communication in chimpanzees and humans, and suggest that both species have unique eye movement strategies for interactions with conspecifics.

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 communication, accompanied by intense eye contact (Argyle & Cook 1976). How and when such unique forms of human communication evolved have long been of interest, particularly from a comparative perspective that addresses how and to what extent human and nonhuman animals are similar and dissimilar in their forms of communication (Chevalier-Skolnikoff 1973; Darwin 1999). Chimpanzees, *Pan troglodytes*, the species closest to humans, are known to share several forms of facial communication 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 (van Lawick-Goodall 1967; Bard et al. 2005). Additionally, the facial musculature of chimpanzees is homologous with that of humans in several respects, and both species show some similar facial expressions (e.g. play faces: Chevalier-Skolnikoff 1973; Parr et al. 2007; Vick et al. 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 et al. 2007) and the holistic (configural) processing of faces (Parr et al. 1998; Tomonaga 1999, 2007b); however, the evidence for holistic face processing in monkeys remains controversial (Tomonaga 1994; Parr & Heintz 2008; Dahl et al. 2009; Adachi et al. 2009).

On the other hand, notable differences between humans and chimpanzees have also been identified. Face-to-face communication in chimpanzees tends to occur in brief spurts and during arousing situations, and the most typical close-range long-bout affiliative communication is 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.
are necessary to reveal both qualitative and quantitative similarities in face scanning. However, known that these species look at the eye regions longer than at main features of faces (i.e. eyes, nose and mouth) intensely. It is also scanned in apes. Humans and monkeys are known to scan the (Keating & Keating 1982; Mendelson et al. 1982; Nahm et al. 1997; Pelphrey et al. 2002; Henderson et al. 2005) and monkeys 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 communication. 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. Therefore, humans might have stronger behavioural tendencies than chimpanzees to collect information actively from the eye regions. Carpenter & Tomasello (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. In this related study, the human–raised and enculturated chimpanzees did not differ from those reared by biological mothers in this regard. Kano & Tomonaga (2009) used the eye-tracking method to measure how humans (adults) and chimpanzees viewed whole-body images of conspecifics and nonconspecifics. 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 nonconspecific 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 previous studies.

In this study we 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 we measured frequencies, durations and probabilities of fixations. Although widely used in studies of humans (Yarbus 1967; Walker-Smith et al. 1977; Pelphrey et al. 2002; Henderson et al. 2005) and monkeys (Keating & Keating 1982; Mendelson et al. 1982; Nahm et al. 1997; Sato & Nakamura 2001; Guo et al. 2003; Gothard et al. 2004), this methodology has not previously been used to investigate face scanning in apes. Humans and monkeys are known to scan the main features of faces (i.e. eyes, nose and mouth) intensely. It is also known that these species look at the eye regions longer than at other facial features and are more likely to start by inspecting the eye regions. The comparisons between humans and monkeys have thus suggested qualitative similarities in face scanning. However, since these previous studies were not particularly designed to compare human and nonhuman primates directly, further studies are necessary to reveal both qualitative and quantitative similarities and differences between the species using a phylogenetically closer species to humans and comparable experimental procedures. This study thus directly compared humans and their closest evolutionary relatives, chimpanzees, using a noninvasive eye-tracking method under unrestrained conditions. The fixation sequences of humans were characterized by sequential fixations over the eyes and mouth, which appeared somewhat to resemble inverted triangular traces (Yarbus 1967; Walker-Smith et al. 1977). The precise comparisons between species enabled us to analyse the characteristic patterns underpinning the sequential/spatiotemporal aspects of eye movements in each species.

We conducted two experiments. Experiment 1 presented both conspecific and nonconspecific faces and examined general similarities and differences in face scanning between chimpanzees and humans. We addressed three questions in experiment 1. The first question was 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 concerned 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 was, which factors contribute 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 et al. 2006; reared by own/other species: Martin-Malivel & Okada 2007) and by phylogenetic relatedness (e.g. conspecific/nonconspecific: 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 (van Hooff 1967; Argyle & Cook 1976). 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 practised 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 differs, 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) to allow comparisons between the studies.
Experiment 1

Participants and apparatus
Six chimpanzees (five females, one male; aged 8–31 years) and 18 humans (11 females, seven males; aged 18–31 years; 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 comprising 14 individuals living in an enriched environment with a 700 m² outdoor compound and an attached indoor residence (Matsuzawa et al. 2006). The outdoor compound was equipped with climbing frames 15 m high, 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 occurred 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; shown in the Supplementary Material). 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. Details of calibration training and procedures are given in Kano & Tomonaga (2009).

Stimuli and procedures
Stimuli consisted of 24 colour still photographs of the faces of 17 species of nonprimate 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 grey 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 (four females, one male) and nine humans (seven females, two 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 in the Results 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 grey 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 that was found in experiment 1, faces were presented for 5 s to humans 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 a session, resulting in no loss of data for either species. To define areas of interest (AOIs), 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 five 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 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 versus 1.3 times, 4.2 versus 0.14 times; human faces: 3.8 versus 2.0 times, 3.7 versus 0.23 times; mammal faces: 4.6 versus 1.5 times, 4.1 versus 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 focused 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 (t5 = 5.27, P < 0.001, t17 = 24.40, P < 0.001, respectively) and of human faces (t5 = 6.70, P < 0.001, t17 = 17.71, P < 0.001, respectively; Fig. 2a). Although this pattern of results did not emerge in chimpanzees with regard to mammal faces (t5 = 1.51, P = 0.19) but did emerge in humans with regard to mammal faces (t2 = 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 focusing on the main facial features revealed a significant interaction between subject species and facial features (chimpanzee faces: F1,2,44 = 11.5, P < 0.001, \( \eta^2_p = 0.34 \); human faces: F1,2,26 = 6.03, P = 0.005, \( \eta^2_p = 0.21 \); mammal faces: F1,22 = 44.7, P < 0.001, \( \eta^2_p = 0.67 \)). Post hoc comparisons between the species revealed that humans fixated on the eye region more frequently than chimpanzees (chimpanzee faces: t2,2 = 2.93, P = 0.008; human faces: t2,2 = 2.73, P = 0.012; mammal faces: t2,2 = 5.30, P < 0.001), and chimpanzees fixated on the mouth region more frequently than humans (chimpanzee faces: t2,2 = 4.21, P < 0.001; human faces: t2,2 = 3.34, P = 0.003; mammal faces: t2,2 = 3.62, P = 0.002; Fig. 2a). Post hoc comparisons within species revealed that, unlike the humans (chimpanzee faces: t2 = 9.29, P < 0.001; human faces: t2 = 7.79, P < 0.001; mammal faces: t2 = 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: t2,2 = 1.44, P = 0.20; human faces: t2,2 = 3.92, P = 0.011; mammal faces: t2,2 = 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: F3,36 = 3.63, P = 0.017; human faces: F2,20,45 = 4.36, P = 0.007, \( \eta^2_p = 0.16 \); mammal faces: F3,36 = 11.68, P = 0.001, \( \eta^2_p = 0.34 \)) and for the mouth (chimpanzee faces: F3,36 = 3.82, P = 0.014, \( \eta^2_p = 0.14 \); human faces: F3,36 = 4.26, P = 0.008, \( \eta^2_p = 0.16 \); mammal faces: F3,36 = 9.13, P < 0.001, \( \eta^2_p = 0.29 \)). 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 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 subsequently to refixate on the eye regions, whereas the chimpanzees were more likely than humans subsequently to 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 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: F2,44 = 1.62, P = 0.20, \( \eta^2_p = 0.069 \); human faces: F2,44 = 0.376, P = 0.19, \( \eta^2_p = 0.072 \); mammal faces: F2,44 = 1.96, P = 0.15, \( \eta^2_p = 0.082 \)). 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, to see whether this had an effect, we allowed the same five human participants to view half of the same human and chimpanzee faces again, but with the presentation lasting 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 and found a significant interaction among the three factors subject species* facial features* facial expressions (chimpanzee faces: F3,36 = 4.14, P = 0.013, \( \eta^2_p = 0.25 \); human faces: F3,36 = 1.62,
We then conducted ANOVAs for eye and mouth regions (subject species * facial expressions) and found that humans viewed the eye regions more frequently than chimpanzees (chimpanzee faces: $F_{1,12} = 8.41$, $P < 0.001$, $\eta^2_p = 0.89$; human faces: $F_{1,12} = 15.2$, $P < 0.001$, $\eta^2_p = 0.94$), and chimpanzees viewed the mouth regions more frequently than humans (chimpanzee faces: $F_{1,12} = 40.1$, $P < 0.001$, $\eta^2_p = 0.96$; human faces: $F_{1,12} = 17.1$, $P < 0.001$, $\eta^2_p = 0.92$). Although ANOVAs (facial features * facial expressions) revealed that chimpanzees differentiated between the facial expressions in terms of the proportion of fixations to eye or mouth regions (chimpanzee faces: $F_{3,12} = 8.37$, $P = 0.003$, $\eta^2_p = 0.67$; human faces: $F_{3,12} = 11.0$, $P = 0.001$, $\eta^2_p = 0.73$), this was not the case for humans (chimpanzee faces: $F_{3,12} = 0.61$, $P = 0.615$, $\eta^2_p = 0.071$; human faces: $F_{3,24} = 0.29$, $P = 0.82$, $\eta^2_p = 0.036$). 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 showed 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 refixations 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 shown by humans, the sequences of fixations shown 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, nonconspecific 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 face 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 the other species. Experiment 2 also demonstrated that although chimpanzees changed their scanning patterns in response 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 showed a more rapid shift in fixation location than 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 communication that use the eyes. Similarly, 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).

Figure 3. The number of fixations as a proportion of the total number of face fixations (mean ± SEM) on each feature of the facial expressions for (a, b) chimpanzees and (c, d) humans. The data were normalized to correct for the differences between facial expressions with regard to the proportions of features.
Irrespective of the selection 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 communication. The persistent and prolonged eye viewing shown 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 advantageous for humans to remain focused on the eye region to recognize these species-specific forms of expression in this 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 children with typical developmental courses (Klin et al. 2002; Pelphrey et al. 2002; Dawson et al. 2004; Dalton et al. 2005). The quick scanning of faces by chimpanzees, on the other hand, might enable them to retrieve overall information (e.g., identity, emotion) from faces coarsely but quickly. As mentioned earlier, compared to humans, face-to-face communication among chimpanzees tends 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 other individuals in that it is more likely that prolonged eye contact functions as a threat signal in nonhuman primates than in humans (Thomsen 1974; Gomez 1996). These findings suggest an intimate connection between face scanning and species-specific forms of facial communication in chimpanzees and humans, rendering their respective eye movement strategies functional in interactions with conspecifics.

The direct causes of 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 caused by the less powerful incentive for chimpanzees, compared with humans, to collect eye information, but might also be caused by 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 nonemotional 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 collect eye information actively. From the ontogenetic perspective, the active viewing of eyes by humans might derive from biologically determined patterns reflecting certain selection pressures in evolution, but might also constitute patterns learned during the course of development. Studies in human infants have shown that infants dramatically increase fixations on the eye region at about 7 weeks of age (Haith et al. 1977). It is known that the differential patterns involving direct/averted gaze emerge early in life (within a few days/weeks) in humans (Farroni et al. 2002), chimpanzees (Myowa-Yamakoshi et al. 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 probably 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 engage frequently 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 interspecies differences in terms of social rank, age, sex, etc. Similarly, previous studies in humans have revealed that East Asians tend to look at the eye regions for shorter durations than Caucasians when presented with photographs of East Asian and Caucasian faces (Blais et al. 2008). Given that the human participants in this study (all Japanese) showed more attention to the eye regions than 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 (Guo et al. 2003; Gothard et al. 2004; Ghazanfar et al. 2006). Of special interest in this regard is the characteristic pattern of human face scanning involving sequential fixations on the eye region. This pattern is consistent with results of previous human studies (Walker-Smith et al. 1977; Althoff & Cohen 1999; Henderson et al. 2005). 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 because of the frequent fixations on the mouth region by the chimpanzees. 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 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 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.

Acknowledgments

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 programmes (D07 and A06). We thank Drs T. Matsuawaza, 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.

**Supplementary Material**

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.anbehav.2009.11.003.

**References**


