



Visual Processing of Faces in Juvenile Western Lowland Gorillas without the Use of Training or Reinforcement: A Pilot Study

Sydney Chertoff^{1,2}, Susan Margulis^{1,3*}, and Jonathan D. Rodgers^{2,4}

¹Department of Animal Behavior, Ecology, and Conservation, Canisius College, USA

²Department of Psychological Sciences, Canisius College

³Department of Biology, Canisius College

⁴Institute for Autism Research, Canisius College

*Corresponding author (Email: margulis@canisius.edu)

Citation – Chertoff, S., Margulis, S., & Rodgers, J. D. (2018). Visual processing of faces in juvenile western lowland gorillas without use of training or reinforcement: A pilot study. *Animal Behavior and Cognition*, 5(3), 292–299. <https://doi.org/10.26451/abc.05.03.04.2018>

Abstract – Cognitive research in zoos is often limited by an inability to consistently train and reinforce subjects. This feasibility study aims to address this problem by examining visual processing of faces by Western lowland gorillas (*Gorilla gorilla*) at a zoo in order to identify gaze patterns without any training or reward. A Tobii Pro X2-30 eye-tracker mounted on a TV monitor was used. Participation by the gorillas was voluntary in that no training or reinforcement was provided. Two gorillas (a four-year-old female and a one-year-old male) attended sufficiently to the screen to collect gaze data. In order to further validate the calibration procedure used with gorillas, 61 human participants were tested using the same equipment and stimuli. Images were unfamiliar gorilla females with neutral expressions. Areas of interest were identified for the eye region, lower face (mouth and nasal areas), and overall face, and the patterns of fixation to each region were examined. Gorillas tended to focus less on the eyes than did humans. The current results highlight the feasibility of this kind of data collection without the use of training or reinforcement.

Keywords – Gorilla, human, Eye-tracking, Visual processing

Very little is known about how nonhuman species process their visual environment. We generally make inferences based on observed behaviors, but the salient features of visual stimuli are difficult to quantify. Much of the early research on visual attention used methods such as gaze duration and direction as an indicator of interest (for a review see Winters, Dubuc, & Higham, 2015). A number of studies have focused specifically on processing of facial images, noting both similarities and differences between human and nonhuman primates (Kano, Call, & Tomonaga, 2012; Kuwahata, Adachi, Fujita, Tomonaga, & Matsuzawa, 2004; Tanaka, 2003). Gaze direction and duration has been quantified for just a few species (primarily *Macaca* and *Pan*) (see for example, Kuwahata et al., 2004; Tomasello, Call, & Hare, 1998). Furthermore, these studies often depended on videotaping and manually scoring gaze direction.

More recently, methodologies have expanded to include the use of eye-tracking technology to more accurately measure visual attention and perception in humans and nonhuman primates. The literature on visual processing in nonhuman primates is substantial. To provide relevant background for the current investigation, we provide a broad explanation of how eye-tracking technology works, and

examples of studies that use eye-tracking systems to address a wide range of questions. Generally, eye-tracking is achieved by collecting patterns of the participant's corneal reflections via infrared light and image sensors. This type of data collection can identify the duration, direction, and pattern of gaze of participants when viewing a variety of stimuli. Whereas our knowledge of visual processing in nonhuman primates, as our closest relatives, is growing, much of the work has focused on a few key species and consequently would benefit from expanded taxonomic breadth.

Chimpanzees (*Pan troglodytes*) and rhesus macaques (*Macaca mulatta*) have been the most frequent participants of eye-tracking studies in nonhuman primates, including ones that analyze how they assess their surroundings. For example, comparisons between chimpanzees and humans demonstrate that both species view the main features of a face in a similar manner with few situational variations (Kano & Tomonaga, 2010). Kano and Tomonaga (2011) also demonstrated that both chimpanzees and humans show higher fixation proportions to specific areas, such as faces, when presented with a broader scene. Furthermore, it has been demonstrated that chimpanzees attend more to social and visual cues from other chimpanzees than from humans, whereas humans respond equally to both species (Hattori, Kano, & Tomonaga, 2010). Tanaka (2003, 2007) found that chimpanzees' apparent preference for images of chimpanzees versus humans could be influenced by social experience, as laboratory chimpanzees tended to prefer images of humans. Rhesus macaques (*Macaca mulatta*) have also been the subjects of comparative perception studies. Taubert and Parr (2011) found that rhesus macaques, chimpanzees, and humans will at times use a holistic processing approach to recognize varying faces. Taubert and Parr found that like humans, chimpanzees can switch to a feature-based processing approach when the face is unfamiliar.

Amongst the great apes, most research has been done on chimpanzees; however, limited comparative data do exist on other species and suggest that chimpanzees, gorillas, orangutans, and humans have a similar pattern of face and eye scanning (Kano et al., 2012). Overall, humans view eyes for longer periods than do gorillas. More recently, Howard, Wagner, Woodward, Ross, and Hopper (2017) used eye-tracking technology to study memory for novel events in both chimpanzees and gorillas, and concluded that there were no significant differences in assessment of stimuli by chimpanzees and gorillas and that both species had improved apparent memory for events that included social models. Howard et al. used a similar methodology to the one we are presenting here; however, their subjects had been exposed to cognitive research and were rewarded following participation. The current study extends the findings of Howard et al. by using similar technology but addressing different questions and using naïve participants.

There has been an increasing emphasis on cognitive research with nonhuman primates in non-laboratory settings such as zoos. Typically, labs have resources and funding to conduct a range of studies, but zoos and sanctuaries do not receive the same type of funding in part because research is not always seen as a priority at these institutions. The lack of funding is one limitation to conducting research in a zoo environment. In addition, research may also be limited by lack of time committed to such efforts, and in some situations the inability of outside researchers to engage in such activities. This is particularly unfortunate, given that cognitive studies can serve as a form of enrichment for animals. By eliminating the need to train the subjects, the amount of keeper involvement decreases and makes the study more accessible to researchers who may not be permitted to interact so closely with nonhuman primates. The potential to work with a greater diversity of species using methods such as those described here, is critical.

Hopper (2017) noted in her review of cognitive research in zoos that sample sizes within any zoo-housed species are very small, usually limited to a few individuals at a single institution. Our research aims to expand on our body of knowledge and address some of the limitations of zoo-based studies. First, we aim to demonstrate the feasibility of conducting a study using eye-tracking equipment with zoo-housed gorillas that have not been trained for cognitive testing. Second, we hope to establish a baseline of gorilla gaze patterns when looking at unfamiliar gorilla faces, compared to the typical patterns seen in humans when looking at unfamiliar faces. This preliminary investigation permits us to demonstrate the feasibility of this approach, and begin to ask comparative questions about visual processing.

Methods

Participants. Six gorillas (one adult male, two adult females, two juvenile females, and one male infant) housed at the Buffalo Zoo had the option to participate in this study. All gorillas are members of one troop housed in a semi-natural enclosure. The care of gorillas adhered to the Association of Zoos and Aquariums (AZA) requirements for welfare. Permission was received from both the institutional Animal Care and Use Committee (IACUC) of Canisius College and the Research Committee of the Buffalo Zoo. Two of the six gorillas (a four-year-old female and a one-year-old male) voluntarily participated. All gorilla data are based on these two individuals. To provide a basis of comparison, 61 humans (all young adults, ages 18-25) were recruited from an introductory psychology class. Students received research participation points for their participation. Data from a separate group of ten human subjects were used to validate use of the default calibration of the eye-tracker. The use of human participants was approved by the Canisius College Institutional Review Board.

Materials. The equipment used was the same for gorillas and humans. A screen-mounted eye-tracker (Tobii Pro X2-30) was affixed to a Samsung UN24H4000 24" (61 cm) 720-pixel LED TV monitor to present the stimulus images. The eye-tracker was able to track the corneal reflection patterns of the participants. The distance between the eye-tracker and gorilla participants varied slightly depending on the lighting conditions, but was generally kept between 50 and 60 cm. This was done to account for variations in seasonal ambient light to maximize the effectiveness of the eye-tracker. For human participants, the screen was kept at approximately 60 cm from the participant. The stimulus set consisted of ten images of adult female gorilla faces with neutral expressions, unfamiliar to the participating gorillas. We chose gorilla females to control for any effects that age or sex of stimulus image might have on gaze patterns. Here we are mainly focusing on feasibility and general comparisons. We chose a single stimulus type to permit an adequate sample size to evaluate, but future studies will utilize a broader array of stimuli. The images included the facial region of the gorillas with natural and semi-naturalistic backgrounds that were blurry compared to the face in the image. Five video clips featuring animated images of insects and music, were interspersed after every two static images. The purpose of the video clips was to try and maintain the interest of the gorillas. Each static stimulus was presented for five seconds. The human subjects observed the same stimulus set as did the gorillas. All stimuli were balanced for brightness and size. Humans and gorillas share a very similar visual system, containing little to no differences in their ability to see colors (Dulai, Bowmaker, Mollon, & Hunt, 1994). Therefore, we did not have to modify the stimuli for species' specific visual systems.

Design and Procedure. The gorillas had no prior experience with cognitive testing. They were first habituated to the presence of the equipment and operator for six weeks before any data collection was attempted. The habituation period utilized the identical equipment setup as the data collection period, differing only in the acquisition of data. During this period, images of non-primate mammals were shown on the screen. Both during habituation and data collection, the operator used a variety of toys and enrichment items to encourage the gorillas to approach. The operator did not directly provide any enrichment, food, or toys to the gorillas to reinforce participation. During the first few days of habituation, oats were scattered on the enclosure floor in an effort to attract the gorillas. Animated movie clips were used intermittently during this time, again to encourage proximity and engage interest.

All habituation and data collection were conducted in the off-exhibit holding area at the Buffalo Zoo, between 8:30-9:00 AM. During this time, the gorillas were in the holding area as part of the normal husbandry routine. The apparatus was positioned in the same place every day, in front of one of the four connected mesh-fronted holding enclosures. The apparatus consisted of a small shelf with a mounted arm on which the screen could be moved up and down depending on the height of the participating gorilla. The screen was generally kept at the height most suited for the juvenile gorillas who chose to participate more readily. The gorillas could move freely amongst the four enclosures. Overhead lighting was maintained at a constant level throughout habituation and data collection. Following habituation, data were collected for 12 weeks, four times per week. Data collection was initiated once a gorilla approached

the apparatus, oriented their face toward the screen, and the software identified their eyes. Default equipment calibration was used for gorillas. Because the gorillas did not have any previous cognitive training, a multi-point calibration procedure was not feasible. Default calibration is the system calibration used by the Tobii software. The equipment recorded the location of the subject's eye gaze as stimuli were presented. These locations were recorded automatically as coordinates on the stimulus screen at 30 gaze coordinates per second. Because the gorillas were free to leave at any time, only data for one or two stimuli were collected at a given time, sometimes resulting in incomplete recordings. The presentation of stimuli was randomized to avoid any biases and to ensure that the gorillas would be exposed to all the images during the limited looking time. Based on the interest of the individual gorilla, the researcher could manually skip the video clips. All sessions were videotaped so that individual animal identification could be verified.

Data from human participants were collected over a period of four weeks in a psychology laboratory on the Canisius College campus. All participants were students in an introductory psychology class who were participating in a related study for which a nine-point calibration was used. The room was illuminated only with a single fluorescent light positioned behind the screen. Consent was obtained from each participant. The only instruction given was to look at the screen. The apparatus was positioned on a stable table and remained in the same spot for each participant. In order to support the use of the default calibration versus the nine-point calibration, we compared these two calibration procedures. For the calibration comparison, ten humans were shown images with an array of five targets placed at the center, above, below, left and right of center on a white screen. Each of these ten participants was instructed to look at the targets as they were presented. Each participant was shown the five targets both following nine-point calibration and default calibration (administered the same way as to the gorillas) in counter-balanced order. Areas of interest were drawn around the targets in the same manner as for the facial stimuli. Mean looking time for nine-point versus default calibration was compared.

Data Analysis. The stimuli were standardized so that the face was relatively the same proportion of the total image for each stimulus. Natural variations in faces are expected to occur and we did not adjust for these. The areas of interest were drawn large enough to account for these variations. We used images that were donated to our research by a photographer to ensure the quality of the images was high enough to be presented on a larger screen. Each image was divided into areas of interest (AOI) to quantify the gaze data. Three AOI's were identified for each stimulus image; face, eyes, and lower face



(mouth/nose). The face AOI was used to identify the overall attention to the stimuli by participants to demonstrate feasibility. The eyes and lower face AOI were used to test the specific hypotheses regarding gaze differences. AOI composition is illustrated in Figure 1. All AOI were drawn slightly larger than the key facial landmarks to account for individual variations in facial characteristics. Across stimulus images, the proportion of the image included in the AOI was generally consistent. The proportion of the stimulus image included in the face AOI ranged from 44.61% to 59.08%, $M = 50.48%$, $SD = 4.60%$. The proportion of the stimulus image included in the eyes AOI ranged from 9.76% to 12.56%, $M = 10.74%$, $SD = 0.84%$. The proportion of the stimulus image included in the lower face AOI ranged from 10.33% to 16.24%, $M = 12.59%$, $SD = 1.86%$.

Figure 1. Example of areas of interest (AOI) for gorilla stimuli. Yellow circle = overall face; green circle = eyes; blue circle = lower face.

A paired *t*-test was used to compare mean looking time when using nine-point versus default calibration. A non-parametric Wilcoxon rank sign test was used to compare differences within subjects with respect to areas of interest. All statistical analyses were conducted in SPSS Statistics Version 24.0.

Results

Comparing nine-point and default calibration, the mean looking time across the five targets showed minimal difference with 20.94 s of gaze on the targets for the nine-point calibration and 19.97 s of gaze on the targets for the default calibration (out of 25 s), there was no significant difference between looking time (paired sample *t* test, $t(9) = 0.52$, $p = .615$, 95% CI [-3.24, 5.18]). This supports the approximate equivalence of these procedures.

Two gorillas stayed in front of the screen long enough to record gaze data. From these two individuals, 24 visual scans were collected (13 for the juvenile female gorilla and 11 for the juvenile male). Figure 2 illustrates example fixations of the four-year-old gorilla female (a), one-year-old gorilla male (b), and a human participant for comparison (c). Both gorillas and humans focused on the face AOI as opposed to the background greater than 85% of the time, indicating that both species successfully viewed the salient areas of the stimuli, demonstrating feasibility of this study.

The human preference for eyes was established in the human data, averaged across stimuli, based on a Wilcoxon Ranked Sign test. Here there was a significant difference between humans looking to the eyes versus all other parts of the stimulus, $Z = -5.685$, $p < .001$. Visual representation of the data (Figure 3) reveals that humans ($M = 64\%$, $SD = 15\%$) focused on the eyes in the stimulus images more than did the gorillas ($M = 36\%$, $SD = 12\%$). In comparing the data from the individual gorilla participants to one another: the four-year-old gorilla ($M = 54\%$, $SD = 33\%$) focused on the eyes in the stimuli more than did the one-year-old gorilla ($M = 23\%$, $SD = 24\%$).

For the four-year-old female gorilla, across her 13 individual gaze samples there was not a significant preference for the eyes versus all other AOIs, based on Wilcoxon Ranked Sign test, $Z = -0.47$, $p = .637$. For the one-year-old male gorilla, across his 11 individual gaze samples there was a significant preference for the other areas of the stimuli versus the eyes, based on Wilcoxon Ranked Sign test, $Z = -2.50$, $p = .012$.

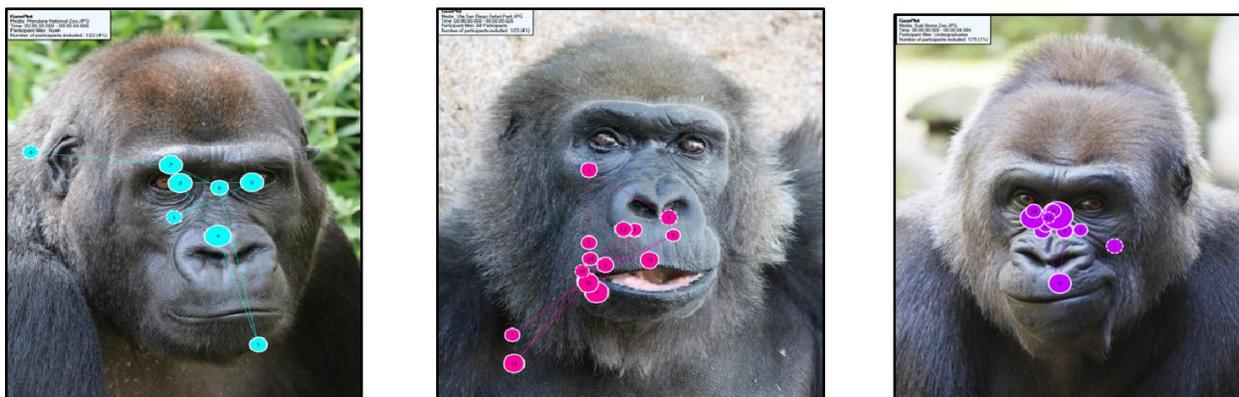


Figure 2. Example fixations of (left) the four-year female old gorilla, (center) one-year-old male gorilla, (right) a human participant.

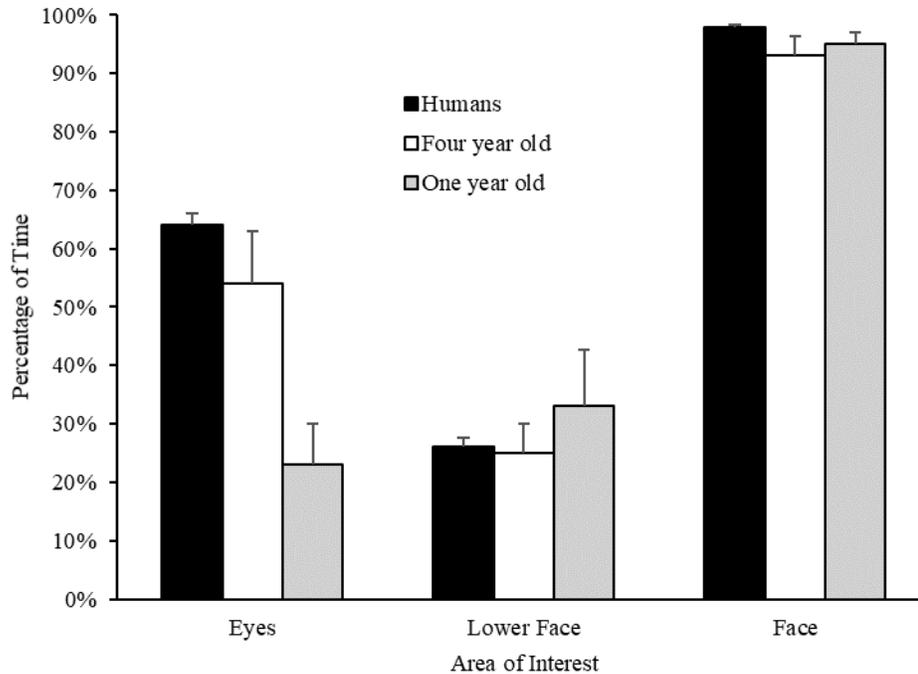


Figure 3. Percentage of fixation time to the eyes, lower face, and overall face region compared between groups. The bars represent the overall averages for human participants (black bar) and the individual gorillas (four-year-old female gorilla, white bar; one-year-old male gorilla, grey bar).

Discussion

Our results demonstrate that it is feasible to conduct a voluntary eye-tracking study in a zoo environment that has not been specifically set up to facilitate cognitive studies, thus satisfying our main aim. Additionally, our results replicate previous findings that humans focused on the eyes more than the gorillas did. We also found that the four-year-old female focused on the eye region more than did the one-year-old male, perhaps suggesting age or sex differences in visual processing.

As a whole, our study demonstrates that eye-tracking equipment can be used successfully with zoo-housed apes, even when they are neither trained to participate nor reinforced for participation. Although our sample size is small, our results nevertheless demonstrate the validity of this approach, and proof of concept for the use of these techniques in a zoo environment. Even though we are optimistic about the potential of this research, it is important to acknowledge that only two of the six gorillas participated. The lack of participation from adult individuals could be attributed to the social dynamics of gorillas and the lack of reinforcement; some individuals may choose not to approach the apparatus when other specific individuals are nearby. In addition, adjusting the apparatus to accommodate different heights of animals can startle gorillas, which further constrains who we were able to collect data from. Prior to, and at points during data collection, we used different enriching items to attract the gorillas' attention. Children's movies appeared to attract the gorillas' attention as did showing the subjects various random items such as colored pens, a stapler, and glow sticks. At one point, all six of the gorillas showed interest; however, it was only the two younger individuals that showed an extended interest and attended to the stimuli long enough for data collection. For example, one of the adult female gorillas would often lie in front of the TV screen and appear to watch the movie, but she would not come close enough to the mesh for the eye-tracker to detect her eyes. To increase participation, we believe that we need to expand

on the stimuli we show the gorillas until we find something that will attract the older individuals for a substantial amount of time. This problem itself poses an interesting question, one we hope to explore in further studies. We intend on using this eye-tracking methodology to present the gorillas with varying items to learn what they show a preference for and then optimize our stimuli based on observed interest. Since the completion of this study, we have successfully obtained eye tracks on two additional gorillas during data collection for a follow-up study.

Although data on human subjects were included only as a basis for comparison, we have preliminary data that suggest that humans focus on the eyes more than young gorillas do. This may be due to information-gathering techniques that vary from species to species. The ability to efficiently gather information from agonistic or affiliative faces is very important for nonhuman primate species (Gothard, Erickson, & Amaral, 2003). Based on this known information about nonhuman primates, it may be that they are better at processing the perceived expression on a face and therefore, do not need to view the eyes as extensively as humans do. Gothard et al. (2003) found that rhesus macaques attend almost entirely to the eyes of a neutral face, whereas they focus on multiple facial features when viewing images of various facial expressions. Another possible explanation for this finding is that eye contact for nonhuman primates, especially great apes, is perceived as a threat and therefore, it would be expected that gorillas would be less likely to look directly at the eyes of the stimulus image (Gomez, 1996).

The results comparing the gorillas as individuals were also interesting. Although the four-year-old female did not look at the eye area significantly more than any other AOI, it must be noted that the eyes comprise a smaller area of the screen than the rest of the image, thus the lack of significant difference may support that she looks more generally to the eyes than other areas of the face. The observed pattern for the four-year-old female looks surprisingly similar to that of humans, as seen in Figure 3. The one-year-old male showed a preference for areas other than the eyes; these differences may be attributable to age, gender, or individual variation, but we cannot distinguish among these possibilities at this time. Previous research shows that male humans and male rhesus macaques pay more attention to threatening faces of conspecifics than to positive stimuli (Lacreuse, Schatz, Strazzullo, King, & Ready, 2013), suggesting that males may be less attentive to non-threatening visual stimuli. An alternative interpretation is based on the ages of the gorillas. Research with young children has shown that perceptual narrowing takes place between six and nine months of age (Pascalis, de Haan, & Nelson, 2002). Six-month-old infants are able to recognize facial identity of both human and nonhuman primates whereas nine-month-old infants and adults are able to recognize only human faces. This finding is useful as it shows that facial processing may change with developmental stage. This idea will be examined in future research.

We continue to explore gaze patterns in gorillas in order to more carefully investigate how gorillas respond to a range of visual stimuli compared to humans and other nonhuman primates that have been studied. Additionally, our ongoing work will enable us to examine developmental changes in visual processing. Our methods open the possibility of using eye-tracking in studies in a greater diversity of species in a zoo setting without the use of training or reinforcement.

Acknowledgements

We would like to thank the Buffalo Zoo for all of their support and help, especially head gorilla keeper Jamie Kranz, general curator Malia Somerville, and President Emeritus Dr. Donna Fernandes. We would also like to thank all of our generous donors without whom we could not have conducted this study, and the team at Experiment.com for their help in raising the funds responsible for this project. S.C. was supported by a General Small Grant from the American Society of Primatologists and the Canisius College Earning Excellence Program. We thank Max Block, Christina Kolbmann, Dr. Michael Noonan, Anthony Syracuse, Jenny Volz, and Anita Yantz for all their beautiful photographs. We are also very grateful to Amy Bucklaew and Elizabeth Tanner for their assistance with data collection.

Ethical Approval

All procedures performed in studies involving animals were in accordance with the ethical standards of the institution or practice at which the studies were conducted. All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from human participants.

References

- Dulai, K. S., Bowmaker, J. K., Mollon, J. D., & Hunt, D. M. (1994). Sequence divergence, polymorphism, and evolution of the middle wave and long wave visual pigment genes of great apes and Old World monkeys. *Vision Research*, *34*, 2483–2491.
- Gomez, J. C. (1996). Ostensive behavior in great apes: The role of eye contact. In A. E. Russon, K. A. Bard, & S. T. Parker, (Eds.), *Reaching into thought: The minds of the great apes*, (pp. 131–151) New York: Cambridge University Press.
- Gothard, K. M., Erickson, C. A., & Amaral, D. G. (2003). How do rhesus monkeys (*Macaca mulatta*) scan faces in a visual paired comparison task? *Animal Cognition*, *7*, 25–36.
- Hattori, Y., Kano, F., & Tomonaga, M. (2010). Differential sensitivity to conspecific and allospecific cues in chimpanzees and humans: A comparative eye-tracking study. *Biology Letters*, *6*, 610–613.
- Hopper, L. M. (2017). Cognitive research in zoos. *Current Opinion in Behavioral Sciences*, *16*, 100–110.
- Howard, L. H., Wagner, K. E., Woodward, A. L., Ross, S. R., & Hopper, L. M. (2017). Social models enhance apes' memory for novel events. *Science Reports*, *7*, 1–7.
- Kano, F., Call, J., & Tomonaga, M. (2012). Face and eye scanning in gorillas (*Gorilla gorilla*), orangutans (*Pongo abelii*), and humans (*Homo sapiens*): Unique eye-viewing patterns in humans among hominids. *Journal of Comparative Psychology*, *126*, 388–398.
- Kano, F. & Tomonaga, M. (2010). Face scanning in chimpanzees and humans: Continuity and discontinuity. *Animal Behaviour*, *79*, 227–235.
- Kano, F., & Tomonaga, M. (2011). Perceptual mechanism underlying gaze guidance in chimpanzees and humans. *Animal Cognition*, *14*, 377–386.
- Kuwahata, H., Adachi, I., Fujita, K., Tomonaga, M., & Matsuzawa, T. (2004). Development of schematic face preference in macaque monkeys. *Behavioural Processes*, *66*, 17–21.
- Lacreuse, A., Schatz, K., Strazzullo, S., King, H. M., & Ready, R. (2013). Attentional biases and memory for emotional stimuli in men and male rhesus monkeys. *Animal Cognition*, *16*, 861–871.
- Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science*, *296*, 1321–1323.
- Tanaka, M. (2003). Visual preference by chimpanzees (*Pan troglodytes*) for photos of primates measured by a free choice-order task: Implication for influence of social experience. *Primates*, *44*, 157–165.
- Tanaka, M. (2007). Development of the visual preference of chimpanzees (*Pan troglodytes*) for photographs of primates: Effect of social experience. *Primates*, *48*, 303–309.
- Taubert, J., & Parr, L. A. (2011). Geometric distortions affect face recognition in chimpanzees (*Pan troglodytes*) and monkeys (*Macaca mulatta*). *Animal Cognition*, *14*, 35–43.
- Tomasello, M., Call, J., & Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. *Animal Behaviour*, *55*, 1063–1069.
- Winters, S., Dubuc, C. & Higham, J. P. (2015). Perspectives: The looking time experimental paradigm in studies of animal visual perception and cognition. *Ethology*, *121*, 625–640.