Development of Joint Attention in Infant Chimpanzees

SANAE OKAMOTO-BARTH and MASAKI TOMONAGA

1 General Introduction

1.1 Gaze Direction, Gaze Following, and Joint Attention

Humans are highly sensitive to the direction of gaze. Determining the precise direction of another's attention is an important ability. Gaze shifts provide salient information about the location of objects but may also function in complex forms of social cognition (Whiten 1997).

In our daily lives, a great deal of information is communicated by means of following another individual's gaze to specific objects and events. This behavioral sequence is called gaze following or joint attention. Gaze following/joint attention is characterized by one individual (X) following the direction of the attention of another individual (Y) attention to an object (Z) (an object of joint focus; Emery 2000). These terms are used interchangeably by researchers. Emery et al. (1997) suggested that gaze following and joint attention are different yet intimately related abilities (probably with different developmental and phylogenetic time courses). They defined gaze following as the ability of X to follow the direction of the gaze of Y to a position in space (not an object). Joint attention has the additional requirement that X follows the direction of Y’s gaze to object Z that is the focus of Y’s attention. Joint attention thus requires extra computation to process the object of attention not just the direction of gaze.

Emery (2000) similarly argued that there are subtle differences between joint attention and shared attention (see also Perrett and Emery 1994). Nevertheless, these two terms are also used interchangeably in the literature. Emery (2000) defined shared attention as a more complex form of communication than joint attention. Shared attention requires that individuals X and Y each have knowledge of the directions of another individual’s attention (or a method for checking that what the other individual is looking at is the same as what they are looking at).
Determining the direction of another individual's attention is easier to establish from more salient visual cues, such as head or body orientation. Thus, attentional cues, or social gaze, are not only provided by the eyes. Perrett and his colleagues have suggested that the direction of the head and the orientation of the body may also provide important indicators of attention when the eyes are obscured or not clear or when the eyes are used for other purposes (Perrett et al. 1992; Perrett and Emery 1994). When all cues are available for processing, a hierarchy of importance exists whereby the eyes provide more important cues than the head and the head is a more important cue than the body.

Research with human infants suggests that sensitivity to gaze shifts occurs very early in infancy. From 3 months of age, infants are able to discriminate changes in an adult's eye direction (Hains and Maier 1996). The ability to follow gaze has been demonstrated most successfully in studies with human infants. In most of those experiments, the general procedure involves the experimenter (or the mother of the infant) sitting face-to-face with the infant. After making eye contact, the experimenter shifts her gaze to a particular location or object. Infant's responses in this task have a specific developmental trajectory. However, the age at which an infant first follows another's gaze is controversial, ranging from 3 to 18 months (Scaife and Bruner 1975; Butterworth and Cochran 1980; Butterworth and Jarrett 1991; Corkum and Moore 1995; D'Entremont et al. 1997). These conflicting results may be due to methodological, conceptual, or definitional differences. Nevertheless, before 12 months of age, human infants follow their mother's gaze but do not direct their attention to the object of her attention. At around 12 months, infants begin to follow their mother's gaze toward particular objects in their visual field, and at around 18 months they can direct their attention to objects outside of their visual field.

Joint attention is considered an early social cognitive ability leading to later developments associated with mental state attribution (e.g., theory of mind, deception, perspective taking; cf. Baron-Cohen 1995; Tomasello 1995, 1999). However, there are some accounts that do not need to attribute understanding other's mental states which apply to young infants. Young infants are primed from an early age to look in the direction that others are looking (cf. D'Entremont et al. 1997). When they do so, they often see interesting (and rewarding) objects and events. Hence, infants may learn to use gaze direction as a cue to where such rewarding events are located (Corkum and Moore 1995, 1998; Moore 1999). In this view, gaze is merely a discriminative stimulus for the general direction in which an attractive event might be encountered, and, once encountered, search should presumably cease. Butterworth's account of gaze following stressed the innate properties of this behavior. He suggests that young infants are hard-wired to follow the gaze direction of others. They are held to terminate search at the first salient object in their scan path (Butterworth and Cochran 1980; Butterworth and Jarrett 1991). In contrast to Moore's conception (1999), there is no learning on the infant's part and no expectation of finding an event.

Although previous studies have suggested that infants are innately sensitive to eye gaze and gaze direction, it is still an open question whether gaze follow-

1.2 Gaze Following and Joint Attention in Nonhuman Primates

Gaze following is also found to occur in a number of nonhuman primates. The use of gaze shifts as social cues has various evolutionary advantages. For instance, gaze shifts may index the location of predators, potential mates, or food sources. Several field studies suggest that primates can follow the gaze of conspecifics (Chance 1987; Menzel and Halpern 1973; Whiten and Byrne 1988). However, in field studies, it is difficult to identify which object or event is looked at two individuals by means of gaze following. For instance, individuals may come to fixate on the same object because the object is inherently interesting even if they do not follow gaze. Such interventional confounds can be effectively excluded in laboratory studies. In fact, several studies have demonstrated that many primate species follow the gaze direction of conspecifics to objects (chimpanzees, mangabeys, macaques; Tomasello et al. 1998; Emery et al. 1997).

The general procedure in primate studies is the same as in studies with humans (see earlier). Furthermore, primates (especially apes) follow the gaze of nonconspecific individuals (e.g., a human experimenter). They do this even when the target is located above and/or behind them (Itakura 1996; Povinelli and Eddy 1996, 1997). Itakura (1996) studied the ability of various species of prosimians, monkeys, and apes to follow a human experimenter's gaze. Only the ape (orangutan and chimpanzee) responded at above-chance levels. Neither Old nor New World Monkeys (i.e., brown lemur, black lemur, squirrel monkey, brown capuchin, whiteface capuchin, stump-tailed macaque, rhesus macaque, pigtailed macaque, and tonkean macaque) responded at above-chance levels.

The clearest evidence for the ability to follow gaze in nonhuman primates comes from laboratory work on great apes, in particular, studies with chimpanzees. Although these studies have shed some light on the topic, they have left many questions unanswered. For instance, how do chimpanzees follow the other's gaze? Which cues are important for gaze following? Why do chimpanzees follow other's gaze? And when do chimpanzees start to follow another's gaze?

Previous studies have investigated how chimpanzees follow another's gaze. Povinelli and Eddy (1996), for example, installed an opaque barrier in a
testing room to obstruct the chimpanzee subjects' line of sight. In cases where the experimenter looked to an object next to the barrier (outside the immediate line of sight of the subject), chimpanzees followed the experimenter's line of sight around the barrier to the unseen object. This ability may be important when trying to extrapolate information from another's attention, specifically, when the focus of attention is out of sight. In another paradigm using a distractor in the subjects' visual field, Emery et al. (1997) reported that monkeys bypassed looking at the first interesting object in their line of sight and followed the demonstrator's gaze to the target object; following gaze geometrically. Tomasello et al. (1999) have reported similar results in chimpanzees. These results suggest that subjects do not reflexively follow gaze to the first available object within their view but actively track the gaze of others geometrically to localize objects or locations others are attending to.

A number of other studies have demonstrated that primates use a variety of cues to track the focus of another's attention (e.g., pointing, head orientation, gazing without head orientation). For example, in one study, Perrett and his colleagues investigated which cue(s) primates use to direct their own attention by measuring eye movements of monkeys during presentation of head and body cues, and head only and head with eyes cues (Lorinz et al. 1999). They found that the subjects used the information from the head more readily than the body. They also appeared to follow gaze cues when the demonstrator's head was oriented toward the subject.

In a series of experiments, Povinelli and his colleagues (1996, 1997) have investigated why chimpanzees follow the experimenter's gaze. The main question is whether they attribute mental states to other individuals when they follow their gaze direction (not automatically). They suggested that chimpanzees can follow an experimenter's gaze but do not use that information to learn about objects in the world or about the "mental state" of the individual providing the gaze cues. As with human infants, it is still an open question whether chimpanzees use gaze as indicators of mental states (as mentioned earlier, however, the question for human infants is at what age do they start to attribute mental states to other individuals' gaze).

Previous studies with chimpanzees, however, have tested adult (or juvenile) subjects and described the ability to follow gaze. A longitudinal study of infant chimpanzees that measures the frequency of gaze following in the course of development may yield important clues as to the ontogeny and evolution of this behavior. Such a study would, in principle, address the question of when (e.g., a developmental time course) infant chimpanzees start to follow another's gaze. A series of longitudinal experiments evaluated the development of joint attention in an infant chimpanzee to address this important question in both human and chimpanzee development.

The present longitudinal series of studies were conducted to clarify the emergence (study 1) and the development (study 2) of the ability in chimpanzees to follow experimenter-given cues, such as gazing and pointing.
2.2 Methods

We tested a male chimpanzee infant, Ayumu, from 6 months to 13 months of age. He had participated in various kinds of cognitive tasks in a face-to-face situation after birth (Matsuzawa 2003; Myowa-Yamakoshi et al. 2003, 2004, 2005). During the experiment, an experimenter faced the infant (Fig. 1) and gave four types of social cues [tap, point, head turn, glance (without head turn)] to one of two objects that were placed in front of the infant (Fig. 2). The subject was given food rewards independently of his responses in the first three conditions, so that his responses to the objects were not influenced by the rewards. In the glance condition, irrespective of the subject's response no food reward was presented. This condition was considered as the test for following the eye-gaze cues. We measured the infant's following responses to the target object from video recordings. The subject also received three kinds of control (non-cued) trials, corresponding to each of the four types of social cues described earlier.

2.3 Results

Figure 3 shows that the infant started to follow the experimenter-given cues (tap, point, and head turn) to the target object in front of him from around 9 months old. By the age of 13 months, the subject showed reliable following responses to the object that was indicated by the glance cue. Furthermore, additional tests clearly showed that the subject's performance was controlled by the "social"

2.4 Discussion

The results of study 1 clearly demonstrate that a chimpanzee infant less than 1 year old can reliably follow the social cues by a human experimenter to shift his attention to a target object. By the age of 13 months, the subject showed reliable following responses to the object that was indicated by the glance cue. This behavior was not controlled by the nonsocial peripheral property (or local enhancement) of the experimenter-given cues.
It has previously been reported that infant chimpanzees less than 3 to 4 years old do not use head-turn cues (Tomasetto et al. 2001). Contrary to this result, our subject reliably used “head turn” as a cue at 11 months and “glance” at 13 months of age. One possible reason for this inconsistency might be differences in the experimental settings. In a study of Tomasello et al., the experimenter looked up to the sky or the ceiling, whereas in the present study, the experimenter looked at the toy as a target object in the subject's visual field. A more important point is that these procedures focus on different aspects of joint attention. As mentioned in the introduction, Butterworth and Jarrett (1991) reported three successively emerging mechanisms of joint attention in human infants from 6 to 18 months. In the present study, we used a specific object within the subject's visual field as a target. This procedure may investigate the “ecological” or “geometric mechanism” of joint attention in their terminology. The Tomasetto et al. study, on the other hand, used no specific objects, but subjects were required to move their heads or bodies. This procedure may require the “representational mechanism.” Tomasetto et al. (1999) reported that chimpanzees follow the gaze direction of other individuals to specific locations geometrically, in much the same way as human infants do. Our experiment used only the two specific locations in the subject's visual field. This procedure might be insufficient for distinguishing between the ecological and geometric mechanisms (see Butterworth and Jarrett 1991). To address problems concerning the underlying mechanisms of gaze following in chimpanzee infants, we need further experimental manipulations, for example, a greater number of specific locations for the target.

In the present experiment, the infant followed the glance cues without explicit differential reinforcement training by 13 months of age. Povinelli and Eddy (1996) reported that 5- to 6-year-old chimpanzees responded appropriately both to head turns and eye movement alone. Itakura and Tanaka (1998) also reported that adult chimpanzees can use eye movement as a cue in object-choice tasks. However, with the exception of the present study, there has been no evidence that infant chimpanzees can use eye movements alone as a cue. Human infants at 9 months of age are unable to shift attention by glance cues without head turning. At 12 and 14 months of age, half the subjects can shift their attention using glance cues (Butterworth and Jarrett 1991; Lempers 1979; but see also Corkum and Moore 1995). These chronological ages apparently correspond to the chimpanzee infant in the present study. However, it is well known that the speed of body growth and perceptual development in chimpanzees is faster than that of humans. We need further studies both in humans and chimpanzees to draw clear conclusions concerning the onset of gaze-following abilities.

As a next step of the investigation of this ability, we might have to focus on the developmental aspect, which includes the “representational mechanism” period of joint attention. In humans, 12-month-olds do not follow gaze to objects behind them but 18-month-olds do (Butterworth and Jarrett 1991). From which age on do infant chimpanzees look back to the target? To clarify this question, we conducted a further study focusing on the spatial representation in joint attention. Moreover, we assessed whether other factors such as the characteristic of target objects affect the infant’s following responses.

3 The Development of Joint Attention

3.1 Introduction

As joint attention has an important role for development in social animals, it is also important to understand how the actions of others elicit infants’ joint attention. Along the lines of Corkum and Moore's empirical and parsimonious account in another study with human subjects, pointing cues elicited more episodes of joint attention than looking alone, and distinctive and complex targets elicited more episodes of joint attention than identical targets (Deák et al. 2000). The authors also found that infants looked more at front than at back targets, but there was also an effect of magnitude of head turn. They also suggested that human infants’ joint attention to targets behind them is affected by the distinctiveness and complexity (i.e., interesting) of the targets. Thus, environmental factors also affect the infant’s joint attention.

Study 2 was conducted to clarify the ability of the infant chimpanzee (Ayumu) to follow experimenter-given cues to targets outside his visual field, the “representational mechanism” in Butterworth's terminology. Moreover, we manipulated two factors to investigate what affects the chimpanzee's joint attention to objects outside his visual field: incentive and subject's memory of targets.

3.2 Methods

In the present study, from 13 months old, Ayumu was tested to look at one of two identical object pairs, which an experimenter indicated by pointing and head turning (Fig. 4). The object pairs were set in front of or behind the subject.

Fig. 4. Ayumu is looking at the target object behind him, which the experimenter is pointing at (from the view of camera 2). (Okamoto et al. 2004)
3.3 Results

Figure 6 shows that the percentages of “looking-back” responses to the distant object by following the distant-pointing cue. The results show that by the age of 20 months the infant reliably began to follow the experimenter’s distant-pointing cue and looked back to the target behind him. Moving targets elicited more responses than stationary targets, and the subject showed more following responses after having seen the experimenter manipulating the computer.

Although Ayuma often turned his head or body to the left side (71.2% of total looking-back responses), the side he turned his head to and the side the experimenter pointed to matched in 80.4% of total looking-back responses. One more important result to be noted is, however, that the subject did not look at the experimenter again once he looked back behind himself.

3.4 Discussion

In the present study, by the age of 21 months, the infant chimpanzee reliably followed the experimenter’s cues and looked back to the target behind him.

Our previous study 1 (Okamoto et al. 2002a) and the present study clearly indicate that mechanisms of joint attention also emerge successively in an infant chimpanzee as in human infants (cf. Butterworth and Cochran 1980; Butterworth and Jarrett 1991). Moreover, factors such as the distinctiveness of targets also influenced the chimpanzee’s joint attention, as in human infants (Dekx et al. 2000). The comparison of results between phase 1 and phase 2 suggests that the subject’s looking-back behavior was facilitated by seeing the experimenter manipulating the computer. Some episodic memory of targets being manipulated may influence joint attention in the sense of increased expectancy of a subsequent event. Furthermore, the attractiveness of subsequent events also affected the subject’s response (moving versus stationary targets, both of which were manipulated by the experimenter; phases 2a versus 2b). In phase 2a, we introduced 2a (moving target) first, and then the 2b (stationary target) conditions alternately. It is possible that the mean responses of 2a are higher than that of 2b because the more interesting condition was presented first and the conditions were not presented in a random order. However, as seen in Fig. 6, the main tendency that the looking-back response is higher in condition 2a than in the adjacent condition 2b is preserved. This tendency indicates that distinctive (e.g., attractive, interesting) targets elicited more looking-back responses than identical ones. Moreover, we stress that the looking-back behavior only occurred when the experimenter indicated the targets, even though the response rate decreased gradually. In other words, only the experimenter’s gesture was used as a trigger to look back to the target behind the chimpanzee infant. Additionally, because the side that the subject looked back to often matched the side the experimenter pointed to, we can suggest that the subject’s responses might represent a “representational mechanism.”
4 General Discussion

4.1 Similarities and Differences of Joint Attention in Chimpanzees and Humans

The present studies found similarities in certain levels of joint attention between human and chimpanzee infants. Two consecutive studies clearly showed that an infant chimpanzee did follow social cues (e.g., tapping, pointing, and head turning) from around 9 months of age. The infant began to reliably follow eye gaze at 13 months. Although, for the gaze cues, a nondifferential reinforcement design was introduced to avoid any type of shaping, it is still possible that the infant may have learned to follow human gaze by “generalizing” from the other cues. Starting at 21 months of age, the infant looked back to targets located behind him, even when there was a distracter in front of him. In this study, infant chimpanzees exhibited the looking-back response, which Butterworth and Jarrett (1991) have interpreted as evidence of a “representational mechanism” in 18-month-old human infants. Although there are some developmental differences about the onset of each “level” of joint attention, on the surface, the development of joint attention in human and chimpanzee infants appears to be highly similar.

However, unlike human infants, the chimpanzee infant in our study failed to look at the experimenter after following her gaze to an object located behind him. This triadic interaction between mother, child, and object of interest has been widely reported in the human developmental literature. For instance, human infants look at the mother and the object alternately. Especially, the situation of study 2 in which the target object was manipulated (movement) might trigger this interaction in triadic relationships between mother and human infant. The absence of this interaction in the infant chimpanzee we observed suggests a fundamental difference in joint attention of humans and chimpanzees.

There might be a potential explanation of the difference of joint attention between the two species. In humans, a number of qualitative changes in social communication occur at around 9 months of age (Carpenter et al. 1998). Human infants, at 6 months, interact dyadically with objects or with a person in a turn-taking (or reciprocally exchanging) sequence. However, they do not interact with the person who is manipulating objects (Tomasello 1999). From 9 months on, they start to engage in “triadic” exchanges with others. Their interactions involve both objects and persons, resulting in the formation of a referential triangle of infant, adult, and object to which they share attention (Rochat 2001; Tomasello 1999). In contrast, Tomonaga et al. (in press) reported that chimpanzee infants never showed an object or gave an object to a caregiver. Such actions are taken to be indicative of referential communication in triadic relationships in humans. One exceptional case of these actions was reported for an 18-month-old nursery-raised chimpanzee by Russell et al. (1997). Yet, unlike human infants who eventually develop an ability to establish this referential communication in triadic...
relationships, it might be possible that chimpanzees never reach this cognitive milestone.

4.2 Prerequisites for Joint Attention

The first appearance of triadic relationship has been treated as an important qualitative change in the development of human infants. The following components are necessary to establish a triadic relationship: infant, adult, and a particular object or event. Previous studies have investigated gaze-following abilities in human infants and nonhuman primates and have used several types of particular objects as targets. Some characteristic of target objects such as distinctiveness or attractiveness elicited the infant’s gaze following behavior in both humans and chimpanzees [Deák et al. 2000; present study 2 (Okamoto et al. 2004)]. These results might be supported by studies that demonstrated that “expectancies” about the appearance of a target elicited more gaze-following responses.

Okamoto-Barth and Kawai (in press) investigated how the expectancy of a target that appeared inconsistently affects attention to gaze-direction cues in human adults. The results showed clear differences between the performances of two groups tested in trials with a consistently and inconsistently appearing target. In trials where the target appeared inconsistently, the subjects disengaged attention when no target appeared because they expected the trial to be a catch trial (no target).

Apes also performed differently in trials where targets appeared consistently or inconsistently. Okamoto-Barth et al. (unpublished data) manipulated the presence of a target in a gaze-following task with all four great ape species (chimpanzees, bonobos, orangutans, and gorillas). When there was no target, the apes showed fewer following responses.

For humans’ daily joint attention episodes, it is possible that a receiver of social attention cues has the expectancy and prediction that a sender may/must be looking at a particular object. Such expectancies and predictions facilitate social communication. However, when there is no particular object or event, the receiver would search for it or wonder what the sender was looking at. In this case, receivers might look back at the sender to gather more contextual information.

A behavior referred to as “checking back” typically accompanies gaze following. This “checking-back” behavior has been observed in adult chimpanzees in instances of apparent uncertainty (Call et al. 1998). However, as mentioned earlier, the chimpanzee infant in the present study did not show this behavior. We interpreted the absence of “checking back” in the infant chimpanzee we studied as evidence against “triad relationship-based” joint attention (or “shared attention”). However, there may be an essential difference between “checking back” and “shared attention.” Checking back could be completed without an understanding of mental states. However, checking back cannot be accomplished without the ability to track gaze geometrically. A number of studies have demonstrated that chimpanzees can effectively track an individual’s gaze geometrically, even when there is a distracter between the subject and the target (e.g., Tomasello et al. 1999). Yet it does not follow from these results that subjects that track gaze are interpreting them mentallyistically. After following the sender’s gaze, does the receiver need a terminal point (target object or event)? Although this is still an open question, it is clear that the presence or expectancy of a target affects the receiver’s attention to a potential target location. For example, consider the preparation for new environmental information that was observed in human adults presented with a schematic face whose eyes were directed to a given location (Okamoto-Barth and Kawai, in press). However, the use of this cue does not imply mental-state attribution in the triadic relationships, but it is more dependent on reflexive orienting.

Is it possible for the chimpanzee to engage in real joint (shared) attention, that is, “triad relationship-based joint attention”? In a dyad interaction, infant chimpanzees do show “mutual gaze” when interacting with their mother from early infancy (cf. Bard et al. 2005; Okamoto et al. 2002b). Furthermore, similar to a report on human infants (Deák et al. 2000), the present studies suggested that some environmental factors influence joint attention to objects outside the visual field of chimpanzee infants. In the future, we should conduct more detailed comparative examinations concerning the developmental changes from dyad- to triad-based interactions involving eye gaze and factors affecting these interactions. Such studies will provide a clearer idea of visual communication including joint attention and the understanding of social-cognitive abilities in nonhuman primates.

Acknowledgments

The present study was supported by grants from the Japan Society of Promotion of Science (JSPS) and the Ministry of Education, Culture, Sports, Science, and Technology (MEXT), Japan (#12020009, #13610086, #16020101), MEXT 21st Century COE Program (A14), and the cooperative research program of the Primate Research Institute (PRI), Kyoto University. Preparation of the manuscript was supported by the Faculty of Psychology, Department of Cognitive Neuroscience, Maastricht University, The Netherlands. Parts of this manuscript were submitted to Maastricht University by Sunae Okamoto-Barth in partial fulfillment of the requirements for a doctorate degree in psychology. We wish to express thanks to M. Tanaka, N. Kawai, K. Ishii, S. Itakura, K.A. Bard, J.R. Anderson, and T. Matsuzawa for their support and advice given in the course of the study. Thanks are also due to J. Barth and F. Subiaul for their suggestions on early versions of this manuscript. We are grateful to the staff of the Center for Human Evolution Modeling Research, PRI, for their daily care of the chimpanzees.
References

Menzel EZ, Halperin S (1975) Paroscopic behavior as a basis for objective communication between chimpanzees. Science 190:652–654
Okamoto-Bart S, Kawai N (in press) The role of attention in the facilitation effect and inhibition of return: Cognition
Tomonaga M, Hare B, Fogalina T (2001) The ontogeny of gaze following in chimpanzees, Pan troglodytes, and rhesus monkeys, Macaca mulatta. Anim Behav 61:335–343