Specificity of gaze-following in young chimpanzees

Daniel J. Povinelli* and Timothy J. Eddy

Laboratory of Comparative Behavioral Biology, University of Southwestern Louisiana, New Iberia Research Center, 4401 W. Admiral Doyle Drive, New Iberia, LA 70560, USA

The development of gaze-following (or joint visual attention) allows human infants to orient to the same object or event to which another person is looking. The development of the ability undergoes elaborations between 6 and 18 months, with older infants displaying the ability to (a) track gaze in response to eye movement alone, (b) look into space outside of their immediate visual field, and (c) precisely localize the exact object at which another is looking. Chimpanzees also exhibit gaze-following, but certain aspects of the specificity of this ability remain unclear. Seven pre-adolescent chimpanzees were used to replicate and extend a previous demonstration of gaze-following and to test the hypothesis that chimpanzees (like 18-month-old human infants) are capable of determining the specific location in space behind them into which another looks. The results strongly support the hypothesis, and suggest extensive commonality in the behavioural manifestations of gaze-following in the two species. Whether chimpanzees develop a parallel understanding of the mental state of attention behind gaze remains less clear.

The development of gaze-following (or joint visual attention) allows human infants to orient themselves to the same object or event at which someone else is looking (Scaife & Bruner, 1975). Although there is some debate about the exact timing of the development of gaze-following in human infants, it may emerge in some form by about 6 months (Butterworth & Cochran, 1980; Butterworth & Jarrett, 1991; Scaife & Bruner, 1975). The ability to extract and use information about the direction of the gaze of another individual is a focal point for theoretical accounts of how and when infants develop an understanding of the mental experiences of others, and in particular the mental state of attention. Some accounts of gaze-following interpret it as an early example of the infant’s awareness of shared attention (e.g. Baron-Cohen, 1994; Baron-Cohen & Swettenham, 1996). On this view, the underlying development of an ability to represent the shared attention of self and others allows for the gaze-following response to emerge, along with other behaviours such as proto-declarative pointing. In contrast, a different account of the emergence of an understanding of joint attention interprets the gaze-following phenomenon as reflecting a kind of ‘hard-wired’ perceptual foundation for the later

* Requests for reprints.

Part of the controversy concerning the proper interpretation of gaze-following relates to developmental changes in specific aspects of the phenomenon. Butterworth and his colleagues have explored the limits of the gaze-following response in human infants of various ages (Butterworth & Cochran, 1980; Butterworth & Jarrett, 1991). Their results indicated that 6-month-olds display some specificity in their gaze-following abilities in the sense that they appear to extract directional information from another's gaze: when an adult shifts his or her head and eyes to the left (or right), so too does the infant. However, at this age at least two limitations in the specificity of gaze-following are evident. First, if the adult’s line-of-sight projects to a point behind the infant, he or she will not turn around to follow the gaze into that space. Thus, 6-month-olds do not appear to represent the adult’s gaze as extending into space which is outside of their own immediate perceptual field—perhaps because they do not represent themselves and the adult as contained within such a space (Butterworth & Jarrett, 1991). Second, if an adult turns to look at a target within the infant’s visual field, 6-month-olds will orient to the first object along their scan path from the adult’s face. However, by 12 months, this latter limitation is overcome and infants appear to possess a ‘geometric’ mechanism for localizing particular objects to which adults glance, although they still will not follow gaze into space behind them (Butterworth & Jarrett, 1991).

Corkum & Moore (1994) have recently provided data indicating that some of the abilities claimed for 6-month-olds may in fact not emerge in a reliable form until about 10-12 months. Nonetheless, there is general agreement that by about 18 months infants appear to have consolidated a number of major aspects of gaze-following. In contrast to younger infants, 18-month-old infants are able to (a) follow an adult’s gaze into space outside of their own visual field, and (b) reliably track gaze in response to joint movements of the head and eyes, or the movements of the eyes alone (but see Corkum & Moore, 1994, for a discussion of additional limitations of 18-month-olds). A number of researchers agree that the sophistication of gaze-following in 18-month-olds may qualify as evidence that they appreciate the aboutness of referential aspect of the gaze of others (e.g. Corkum & Moore, 1994). Consistent with this interpretation, other researchers have shown that 18-month-olds appear to use gaze orientation to associate a nonsense word with a novel object and to discover the reference of an affective outburst (Baldwin & Moses, 1994).

As part of a broader effort to determine when various aspects of social understanding evolved, Povinelli & Eddy (1996a) explored aspects of the gaze-following response in chimpanzees. Their research focused on determining if chimpanzees reliably track the gaze of others, and if so whether they appreciate gaze as a projection of attention. Their results suggested that chimpanzees may share some of the more sophisticated aspects of gaze-following typical of 18-month-old human infants. For example, the chimpanzees selectively looked behind themselves after witnessing a human experimenter shift his or her gaze from their faces to these areas, and they did so both in response to movements of the head and eyes in concert, as well as the eyes alone. The purpose of the present study was to replicate the results of Povinelli & Eddy (1996a) in a different context, as well as address certain limitations of their study. First, although their subjects reliably glanced somewhere behind themselves when the experimenter did so, as compared to control
Gaze-following in chimpanzees

trials where the experimenter did not glance at all, the specificity of this gaze was less clear. Although the subjects tended to look to the left (or right) when the experimenter did so, the effect only approached statistical reliability. Thus, their study did not rule out the possibility that chimpanzees might possess an analogue of the human gaze-following response that does not include all aspects of its specificity. For example, they might possess or learn rules about gaze information which allows them to search space behind them in response to the gaze of others, but might not deploy their response as selectively as 12- or 18-month-old human infants. Thus, using additional controls not present in the Povinelli & Eddy (1996a) study, we sought to simultaneously test two hypotheses about the specificity of the chimpanzees’ gaze-following response. First, we attempted to replicate the finding that chimpanzees can use another’s gaze to search into space outside their own immediate visual field, and second we tested the hypothesis that chimpanzees are capable of extracting information concerning the specific space behind them indicated by an experimenter’s gaze.

Method

Subjects

The subjects were seven captive-born pre-adolescent chimpanzees living at the University of Southwestern Louisiana New Iberia Research Center. At the time the study began, the subjects ranged in age from 5 years 9 months (5:9) to 6:7. Six were female (Kara, Jadine, Brandy, Megan, Mindy, Candy) and one was male (Apollo). All of the animals had been living in social groups with other chimpanzees their entire lives. At the time the present study was conducted, they had been living together in a facility consisting of five indoor-outdoor housing units which were interconnected by a series of shuttle doorways which could be closed off as necessary. They had free access to all of these areas except during testing (see below). They were part of a long-term research programme in comparative cognition and had participated in a number of studies of their cognitive abilities, including their capacity to recognize themselves in mirrors, and their understanding of mental states. A more detailed description of their rearing and testing histories is provided in Povinelli & Eddy (1996b).

Procedure

General setting. The experiment occurred in two of the outdoor enclosures of the subjects’ living area (‘waiting’ and ‘testing areas’, Fig. 1). Before each training or testing session, the subjects were ushered out of the areas to be used and a shuttle door was closed to prevent them from entering these areas. Next, they were individually transferred out of the group by having a trainer open the shuttle door and allowing the subject to enter the waiting area. All of the subjects were intimately familiar with this procedure of transferring out of the group for individual testing. After the subject was isolated from the group, he or she could be transferred between the waiting and testing areas by having the trainer open the shuttle door (via a pulley system located outside the enclosures) which connected the two areas. Before any training or testing began, two video-cameras were set up outside the testing area in the positions depicted in Fig. 1. One of the cameras was in front of the extreme right hand side of the enclosure, and the other was on the extreme left hand side. Both were oriented to the centre of the enclosure where an experimenter sat (see below) and each provided a view of most of the enclosure.

Training procedure. A trainer and an experimenter participated in all trials. The role of the trainer was to open the shuttle door connecting the waiting and testing areas at the beginning and end of each trial, allowing the subject to enter and exit the test area. The experimenter sat inside the testing area as shown in Fig. 1, and provided food rewards as described below. In order to provide a standardized context in
which we could later deliver the experimental treatments (see below), the subjects were each trained to execute a simple two-step routine when the trainer initially opened the shuttle door. The first step was to enter the testing area where the experimenter sat, approach him along the predetermined pathway indicated in Fig. 1, and take a food reward from him. The second step was to turn around and return to the waiting area by climbing through the shuttle door. With this routine established, the experimenter would be able to execute the required treatments while the animals were in a standardized location in space. On each training trial, the experimenter made a judgment as to whether the subject had successfully completed the two actions, and then relayed this information to the trainer (who recorded it) at the end of each trial. If the subject did not approach the experimenter along the specified pathway, the animal was ushered out of the test area into the waiting area and was not rewarded. In an attempt to minimize potential distractions caused by the presence of the trainer, after initially opening the door to allow the subject into the testing area, the trainer stepped behind a solid wall (see Fig. 1) and remained out of sight as the chimpanzee responded to the experimenter. Each subject received three trials per

Figure 1. Setting and procedure for training and testing chimpanzees. On experimental treatment trials, once the chimpanzees reached the starting point (depicted by head of arrow), an experimenter shifted his gaze to either the left or right camera lens. Shaded areas indicate regions into which the chimpanzees' glances were coded (see text for details).
Gaze-following in chimpanzees

training session and typically received one session per day. In order to advance to testing, the subjects were required to meet a formal criterion of two sessions of flawless performance on both components of the entering and exiting routine.

**Testing procedure.** The testing phase of the experiment occurred across 32 sessions, with two sessions administered back-to-back to each subject on each day. Thus, for each subject the experiment lasted 16 days. One of the subjects' two daily sessions consisted of four trials which were identical to the training trials. These are referred to as 'no treatment' sessions. The other daily session again consisted of four trials; three of these were identical to those used in training. On the remaining trial, the experimenter executed one of four treatments. These are referred to as 'treatment' sessions. The two experimental treatments were glance-right (GR) and glance-left (GL) trials. These trials were implemented by having the experimenter follow the chimpanzee with his eyes until the animal reached the starting point (see Fig. 1). Once the subject had reached this point, the experimenter pivoted his head and upper torso and looked directly into either the left or right camera, depending on the scheduled treatment (see below). This posture was held for exactly 5 s, after which the experimenter always rewarded the subject, regardless of the animal's behaviour during this time period. In order to execute the treatments for the precise duration of 5 s specified by the protocol, the experimenter wore an earplug attached by a cord (obscured under clothing) to a hand-held timer which was set to beep 5 s after activation. (The experimenter had worn the earplug and held the timer during the training phase in order to accustom the subjects to its presence.) Thus, on each treatment trial, the experimenter watched the subject cross through the shuttle door and as the subject reached the starting area he began the timer and executed the scheduled treatment. After the timer signalled the end of the 5 s period, the experimenter ceased executing the treatment, praised the subject and handed the subject a food reward. Finally, the trainer opened the shuttle door, allowing the subject to return to the waiting area.

In addition to the two experimental treatments (GR, GL) described above, two control treatments were also used. For no-glance (NG) control trials, after the subject reached the starting point the experimenter simply looked at the subject for 5 s before handing the subject its food reward. Just before executing this treatment, the trainer engaged in a slight downward (nodding) movement of the head. This motion was necessary so that raters who would later code the tapes of these trials would have a visual cue for when the 5 s window for the trial had begun. Finally, for irrelevant-motion (IM) control trials, the experimenter moved his right leg and thigh back and forth (for 5 s) so that his right foot slid along the floor of the enclosure.

The order of no treatment and treatment sessions on each day were counterbalanced so that each type appeared equally often on the first and second session on a given day across subjects. Within that constraint, the order of the treatment and no treatment sessions was randomized for each subject. Assignment of the experimental and control treatments occurred such that each subject received four trials of each of the four treatments (GR, GL, NG, IM) across the course of the experiment. With one treatment trial per session, a total of 16 treatment sessions were required (4 treatments × 4 trials = 16 treatment sessions).

The order in which the 16 treatment trials were administered to each subject was determined as follows. First, the subjects were randomly divided into two groups (Group 1 = Kara, Brandy, Apollo, Candy; Group 2 = Jadine, Megan, Mindy). Group 1's first treatment session involved experimental treatments (either GR or GL as randomly assigned), whereas Group 2's first treatment session involved control treatments (NG or IM as randomly assigned). Thereafter, the two groups alternated between treatment sessions that contained experimental treatments and those that contained control treatments. Within that constraint, the 16 treatment trials for each subject were randomly and exhaustively assigned. Thus, on any given day, roughly half of the subjects received experimental treatments in their treatment session and the other half received control treatments. The specific treatments were also counterbalanced so that each one appeared equally often when the treatment session was the first session of the day as when it was the second session. Finally, the trial number on which the experimental and control treatments were executed was randomly assigned. (As indicated above, treatments were initiated when the subjects reached the starting point on their path of approach to the trainer. However, in some cases the subjects deviated their approach to the trainer in such a fashion that they did not cross the starting point. On these trials, the experimenter had been instructed not to execute the treatment and it was automatically reassigned to the following trial. In rare cases it was necessary to reassign the treatment to
the next trial as well. Reassignments were necessary on 26.8 per cent of all treatment trials, with an approximately equal frequency for GR and GL trials.)

Videotape codings. A set of standardized instructions and summary sheets were given to four undergraduate students who had not participated in the study. The instructions briefly described the general procedure of the experiment without specifying the purpose of the treatments executed by the experimenter (although his actions were clearly visible on the tape). The observers were instructed to view the two tapes, the one recorded from the left camera and the one recorded from the right camera. The instructions further requested the raters to view each treatment trial as recorded from each camera (right and left) for each animal. They were instructed to start a timer when they observed the experimenter initiate the treatment, and then to observe the animals' behaviour for 5 s. For each trial, a judgment (yes or no) was requested for whether 'the subject looked behind itself in the general direction of the camera which made the tape you are viewing'. These verbal instructions were augmented with a diagram of the enclosure similar to Fig. 1 which indicated the areas that constituted 'the general direction' of the two cameras (see shaded regions of Fig. 1). Thus, for each trial, each rater made a judgment about whether the animal looked behind itself in the direction of the left camera, the right camera, both, or neither.

Inter-observer reliability. Each of the four raters coded a total of 224 probe trials (7 animals × 16 probe trials × 2 cameras). Inter-observer reliability scores were calculated for each of the six possible pairwise combinations of observers by dividing the total number of agreements by the total number of opportunities for agreement. Overall reliability across the six pairwise comparisons averaged 80.5 per cent.

Data summary and analysis. Probe trials for each animal were first summarized for whether the observers indicated a positive response (regardless of frequency') for glancing in the general direction of the camera. In order to provide the most accurate measure of the subjects' behaviour possible, for each rater and camera (right or left) the percentage of trials (out of 4) that were judged to contain a glance were calculated for each of the four treatments. Next, the separate scores of each of the four raters were averaged to produce a mean percentage of trials containing glances for each subject for each treatment type (for both the right and left cameras). This allowed us to examine the specificity with which the subjects glanced in the same direction as the experimenter (GL and GR treatments), as well as the ambient level of glances directed at each camera (as operationalized by the NG and IM control treatments). The data were analysed using a one-way repeated measures analysis of variance (ANOVA) with the data arranged so that each subject contributed a single score to each of the four treatment columns.

Results

In terms of the main hypothesis, the results of the two experimental treatments (GL and GR) can be understood as matching or mismatching glances to each camera. Matches represent those occasions on which the subjects glanced in the same direction as the experimenter, whereas mismatches represent those occasions on which the subjects glanced in the opposite direction as the experimenter. The main results of the experiment are depicted in Fig. 2. The top panel of the figure displays the average percentage of trials on which the subjects glanced at least once in the direction of the right camera during each of the four treatments. The bottom panel depicts the same data for the left camera. Thus, when the experimenter glanced to the right camera (Fig. 2, top panel), the subjects glanced behind themselves in the same direction (matches) on 49.6 per cent of the trials.

1 Frequency data were also collected and analysed. The patterns obtained by these data were statistically indistinguishable from the patterns present in the nominal data reported here due to the low frequency of multiple glances by subjects within the 5 s trial duration.
In striking contrast, when the experimenter glanced to the left camera, the subjects exhibited glances to the right (mismatches) on only 4.4 per cent of the trials. The NG and IM controls for the subjects’ glances to the right were low and nearly identical to levels of mismatched glances to the right. A one-way repeated measures ANOVA compared the percentage of glances in each of the four treatments and revealed a strong overall effect \(F(3,18) = 36.538, p < .001\). Tukey-Kramer Multiple Comparison post-tests further indicated that as predicted the percentage of GR trials containing glances to the right was significantly greater than the percentage of GL, NG or IM trials containing glances to the right \(p < .05\) in all three cases). Furthermore, the post-tests indicated that the percentage of trials containing glances to the right did not differ among the GL, GR or IM trials.

![Figure 2](image-url)

**Figure 2.** Percentage of trials during which chimpanzees glanced behind themselves to the right or left as a function of where (right or left: GR and GL, respectively) or if (no glance or irrelevant motion: NG and IM, respectively) an experimenter looked. See text for detailed description of treatments.

The same pattern is evident for the subjects’ glances in the direction of the left camera (Fig. 2, bottom panel). When the experimenter glanced to the left camera the subjects displayed matching glances behind themselves to the left on 80.4 per cent of the trials. When the experimenter glanced to the right camera the subjects displayed mismatched glances to the left on 47.0 per cent of the trials. An examination of the NG and IM controls for the subjects’ glances to the left reveals that the subjects showed a much higher frequency of ambient glances to the left than they showed to the right (as is
obvious by comparing the NG and IM controls in the top and bottom panels of Fig. 2). This is probably because the left portion of the testing area was where the subjects entered and exited the testing area, and the subjects therefore may have been monitoring the shuttle door or waiting for the trainer to reappear. Nonetheless, the ambient level of glances to the left during the NG and IM controls was comparable to the frequency of mismatched glances to the left, and far lower than the matched glances to the left. A one-way repeated measures ANOVA was conducted in order to compare the percentage of glances in each of the four treatments, and revealed a strong overall effect ($F(3,18) = 10.821, p < .001$). Tukey–Kramer Multiple Comparison post-tests indicated that the percentage of GL trials containing glances to the left was significantly greater than the percentage of GR, NG or IM trials containing glances to the left ($p < .05$ in all three cases). As with the data for the right camera, the post-tests further indicated that the percentage of trials containing glances to the left did not differ among the GL, GR or IM trials. Thus, despite the fact that the mismatched glances to the left during the GR treatment were considerably higher than the comparable data for the right camera, this elevation was due a uniformly higher frequency of glances to the left camera during all treatments (as compared to the right camera).

Finally, in order to satisfy the concerns of a reviewer, unplanned comparisons within individual raters for the group were conducted. With respect to the data for matching glances to the right camera, the results of three raters yielded the predicted pattern (GR highest), and one-way repeated measures ANOVAs indicated that this effect for three of the four were significant at the $p < .05$ level. Similarly, with respect to the data for matching glances to the left camera, all four raters showed the predicted pattern (GL highest), and two were significant at the $p < .001$ level. Finally, if the comparisons are restricted to GR and GL alone, every group comparison for every rater was in the direction of the pattern predicted.

**Discussion**

Both of the hypotheses we tested were supported. Replicating the findings of Povinelli & Eddy (1996a, b), the chimpanzees responded to an experimenter who glanced behind them by turning to look behind themselves. In addition, the quadrant of space into which the chimpanzees looked matched the space into which the human glanced at levels far exceeding chance. Furthermore, the results of the control trials mitigate the possibility that some general aspects of movement on the part of the experimenter aroused the subjects, which consequently increased their overall levels of visual scanning. The subjects’ glances in the direction of the target camera were significantly lower on those trials where the experimenter engaged in an irrelevant motion (or no motion at all) for 5 s, as compared to those trials (GR, GL) where the experimenter glanced to the target camera for the same duration.

In combination with previous results, we believe the following conclusions can be drawn about the specificity of gaze-following in chimpanzees. First, the results of the present study demonstrate that 6- to 7-year-old chimpanzees (and quite possibly even younger chimpanzees) extract specific information about the direction of gaze from others, at least when the gaze occurs in concert with head and upper torso movement. Second, chimpanzees display the gaze-following response whether it is instantiated by
movements of the head and eyes in concert, or simply the eyes alone (Povinelli & Eddy, 1996a). Third, chimpanzees will use another’s gaze to visually search into space outside their immediate visual field in response to eye plus head/upper torso movement, eye plus head movement, or just eye movement alone (Povinelli & Eddy 1996a). Indeed, Povinelli & Eddy (1996b) have shown that chimpanzees do not even need to witness the shift in another’s gaze direction in order to follow it into space outside their immediate visual field. Rather, the gaze-following response can be triggered when chimpanzees encounter another’s head and eye direction statically oriented above and behind them. Finally, in addition to these conclusions, the results of Povinelli & Eddy (1996a) suggest that chimpanzees may also possess some understanding of how another’s gaze is interrupted by solid, opaque surfaces.

There are at least two possible interpretations about the level of social understanding in chimpanzees based on what is currently known about the specificity of their gaze-following abilities. One possible interpretation of the striking commonality in gaze-following in humans and chimpanzees is that it reflects a similar mentalistic understanding of attention in both species. That is, given that human infants of 18 months appear to extract genuinely mentalistic information about the attention of others from their head posture and gaze direction, and given that the sophisticated aspects of gaze-following appear to emerge at about this same age, one might conclude that the chimpanzee’s sophisticated gaze-following behaviour reflects a similar kind of appreciation of attention. However, in an extensive series of 14 studies with these same animals, Povinelli & Eddy (1996b) found little or no evidence that they understood seeing as a mental event. For example, when given a choice between two experimenters, the animals were as likely to use their visually based begging gestures to the experimenter who could not see them (i.e. because he or she wore blindfolds) as to the experimenter who could see them.

Given the possible decoupling of an explicit awareness of the mental state of attention from gaze-following, a second possibility is that gaze-following, including its discrete developmental elaborations, may have evolved in organisms who did not possess an understanding of each other as agents with internal, attentional states. For example, Povinelli & Eddy (1996a) have speculated that gaze-following may have evolved as a kind of commensal relation in which useful information about predators (e.g. van Schaik, van Noordwijk, Warsono & Sutrisno, 1983) or social exchanges (e.g. Chance, 1967) could be gathered at no cost to either individual involved. With respect to its development in human infancy, Butterworth & Jarrett (1991) and Corkum & Moore (1994) have similarly argued that perhaps the origins of gaze-following can be found in fairly low-level psychological phenomena, but that this initial mechanism serves to draw infants into a world of shared meaning from which they ultimately construct the notion of the mental state of attention. In addition, our research with chimpanzees further cautions that other species may display the same degree of specificity in the response as 18-month-olds, without necessarily understanding the mental state of attention per se. Thus, gaze-following in human infancy—even in 18-month-olds—may have little to do with an underlying appreciation of attention. Rather, because of other conceptual developments occurring in parallel, it may be that at some point in their development humans reinterpret gaze-following with a subjective meaning that it did not originally possess.
Acknowledgements
This research was supported by NIH Grant No. RR-03583-05 to the University of Southwestern Louisiana New Iberia Research Center and NSF Young Investigator Award SBR-8458111 to DJP. Dr William E. Greer provided generous access to subjects and facilities. We thank Anthony Rideaux for training the animals and Walter Welch for participating in testing, Bridget Beaullieu, Elicia Lacey, Cindy Minnaert and Laura Theall assisted in videotape coding and/or data analysis. Original artwork is by Donna T. Bierschwale.

References

Received 22 September 1995; revised version received 1 April 1996