Articles

Joint Attention and Imitative Learning in Children, Chimpanzees, and Enculturated Chimpanzees*

Malinda Carpenter and Michael Tomasello, Emory University and
Sue Savage-Rumbaugh, Georgia State University

Abstract

In this study we compared the nature of the joint attentional interactions that occurred as chimpanzees and human children engaged with a human experimenter (E). Subjects were three chimpanzees raised mostly with conspecifics (mother-reared), three chimpanzees raised in a human-like cultural environment (enculturated), and six 18-month-old human children. Of particular interest were possible differences between the two groups of chimpanzees that might have resulted from their different ontogenetic histories. Observations were made as subjects participated in an imitative learning task involving a number of novel objects. Variables coded were such things as subjects' looks to the object, looks to E, the coordination of such looks in periods of joint engagement with E, and gestural attempts to direct E's attention or behavior (declaratives and imperatives). Results showed that enculturated chimpanzees were most similar to human children in social interactions involving objects, for example, in their attention to the object in compliance with E's request, their joint attentional interactions during less structured periods, and their use of declarative gestures to direct E's attention to objects. They were not similar to children, but rather resembled their mother-reared conspecifics, in the duration of their looks to E's face. A positive relation between subjects' joint attentional skills and their imitative learning skills was found for both chimpanzee and human subjects. It is concluded that a human-like sociocultural environment is an essential component in the development of human-like social-cognitive and joint attentional skills for chimpanzees, and perhaps for human beings as well.

---

* This investigation was supported in part by National Institute of Child Health and Human Development grant NICHD-06016, which supports Georgia State University's Language Research Center and by RR-00165 from the National Center for Research Resources to the Yerkes Regional Primate Research Center. The Yerkes Center is fully accredited by the American Association of Laboratory Animal Care. This research was also supported by the College of Arts and Sciences of Georgia State University. Thanks to Rose Sevcik for help with data collection, to Jennifer DeLong for help with coding, and to Gerry Lawrence and Josep Call for help with data analysis. Address correspondence to Malinda Carpenter, Department of Psychology, University of Denver, CO 80208, USA.

© Blackwell Publishers Ltd. 1995. Published by Blackwell Publishers, 108 Cowley Road, Oxford OX4 1JF, UK and 238 Main Street, Cambridge, MA 02142, USA.

Keywords: Communication; imitation; joint attention; primates; social cognition; social learning.

At around 12 months of age human infants begin to engage with adults in relatively extended bouts of joint attention to objects (Bakeman & Adamson, 1984). In these interactions infants actively coordinate their visual attention to person and object, for example, by looking to an adult periodically as the two of them play together with a toy. Infants also become capable at this age of intentionally communicating to adults their desire to obtain an object or to share attention to an object – usually through nonlinguistic gestures such as pointing or showing, often accompanied by gaze alternation between object and person (Bates, 1976). It is also at around this same age that infants first imitatively learn from adults novel actions in which they reproduce not only a demonstrated change of state in an object but also the behavioral strategy of the adult demonstrator (Meltzoff, 1988a; 1988b). The emergence of these various 'triadic' skills at around the same age is not an accident, we would argue, but rather a reflection of their common reliance on infants' emerging ability to understand other persons as intentional agents whose attention and behavior to objects may be actively followed into, directed, and shared (Tomasello, 1995).

An important question is whether these social-cognitive skills of human infants – sometimes referred to as 'secondary intersubjectivity' (Trevarthen, 1979) – are uniquely human skills. Of special importance for answering this question is our nearest primate relative, the chimpanzee (including both *Pan troglodytes* and *Pan paniscus*). In their natural habitats chimpanzees show some behaviors seeming to indicate skills of secondary intersubjectivity, but there are also some important differences with humans. For example, in their natural habitats chimpanzees routinely follow the visual gaze of others to interesting and useful objects and events (Plooij, 1978), but it is not clear that they engage in more extended bouts of joint attention to objects (at least not as defined by human researchers). They also use a variety of means of intentional communication to directly affect the behavior of others in the wild (imperatives), but again it is not clear (there are only a few anecdotal reports) that they make concerted attempts to direct the *attention* of conspecifics to outside objects (declaratives), and they clearly do not use the prototypically human types of attention direction in the form of pointing and showing (Plooij, 1978; Tomasello, 1990). Finally, although chimpanzees are clearly capable of various forms of social learning from conspecifics, they do not seem to engage in the imitative learning of actions on objects in which they reproduce both the end and means of a novel behavior or behavioral strategy (Tomasello, 1994).

These differences between chimpanzees and humans may be based on biological differences in the two species' capacities for entering into certain kinds of joint attentional engagement with conspecifics. But it is also possible that they are based, at least in part, on differences in the environments in which the two species develop. Most important for current purposes, objects, and especially artifacts, do not play as important a role in the lives of developing chimpanzees in their natural habitats as they do in the lives of human children. This is true both when they behave in isolation as well as in social interactions. Thus, adult chimpanzees do not routinely point to, show, or give objects to their offspring; instruct them in their use; or integrate objects into their social interactions on a regular basis. In

© Blackwell Publishers Ltd 1995

over a decade of observation, Boesch (1991) reports just two examples of direct behavioral instruction with objects (and these have other interpretations), and indeed one of the most striking findings in the cross-species comparison of Bard and Vauclair (1984) was the relative absence of attempts by chimpanzee adults to instruct or direct the attention of youngsters to objects.

The question thus arises whether chimpanzees raised from an early age in more human-like cultural environments – with more exposure to objects and artifacts, along with humans who direct their attention to objects and instruct them in their use – develop more human-like skills of secondary intersubjectivity. Informal observations would seem to suggest that they do. It is a common occurrence, for example, that enculturated chimpanzees and their human caregivers engage in joint attentional interactions around objects with alternating visual attention between partner and object (Savage-Rumbaugh, 1990). Intentional communication about objects is also quite common, with some use of protodeclarative pointing and showing (but cf. Gómez, 1991; Gómez, Sarría, & Tamarit, 1993). Imitative learning would also seem to be in the repertoire of enculturated chimpanzees, as some have seemingly acquired their symbolic communication skills solely by means of observation (Savage-Rumbaugh, McDonald, Sevcik, Hopkins, & Rubert, 1986). The only study to make a direct comparison among chimpanzees, enculturated chimpanzees, and human children is reported by Tomasello, Savage-Rumbaugh, and Kruger (1993), who found that enculturated chimpanzees imitatively learned both the end and means of novel actions on objects just as well as 2-year-old human children and much better than their mother-reared conspecifics (see also Hayes & Hayes, 1952).

In the current study we were interested in seeing whether a similar pattern might be found in the joint attention and gestural communication displayed by these same three groups of subjects. We thus re-coded the videotapes of the Tomasello, Savage-Rumbaugh, and Kruger (1993) imitation study, this time on a second-by-second basis with a focus on the looking patterns of chimpanzee, enculturated chimpanzee, and child subjects as they interacted with the human experimenter. We also noted any attempts made at gestural communication. Our interest was both in species differences – chimpanzees compared to human children – and in environmental differences between the two groups of chimpanzees who had been raised in different ways – mother-reared or enculturated by humans. We expected that the human children might show a species advantage in such things as the average length of joint attentional episodes and the average length of looks to the experimenter's face, due to differences in the typical pattern of attention deployment of the two species and/or to the different social motivations of the two species in interacting with a human experimenter. We also expected that the enculturated chimpanzees would show an advantage over mother-reared chimpanzees in such things as the average length of their joint attentional episodes with humans and in using declarative gestures to indicate an interest in sharing attention to objects with humans – quite simply because in both of these types of interaction with humans enculturated chimpanzees have had much more experience.

A second goal of the current study was to determine what, if any, relations exist between joint attention and imitative learning for all subjects. These two skills are thought by some researchers to be related to one another because the imitative learning of an action on an object requires that the learner pay attention

© Blackwell Publishers Ltd. 1995

to and reproduce not only the end result effected on the object, but also the behavioral means of the demonstrator (Tomasello, 1990, 1995). This might require, in some instances at least, a coordination of looks between the object whose state is changing and the person whose behavior is changing its state. To test this possibility, we attempted to relate the performance of individual subjects on the imitation tasks with their joint attentional looking patterns as they engaged in the tasks. In general, it was expected that individuals who engaged in more and longer bouts of joint attentional engagement also would show more sophisticated skills of imitative learning. We had no theoretical expectations that the nature of this relationship should be different for the different groups of subjects.

Method

Subjects

Six chimpanzees from the Language Research Center of Georgia State University served as subjects (Ss). Three of these – the enculturated chimpanzees – were raised from infancy in an environment much like that of human children. They interacted daily with human caretakers in an environment full of objects and artifacts and received various types of instruction and attention encouragement with these objects and artifacts. They were also exposed from an early age to a system of symbolic communication involving a lexigram keyboard (and to spoken English as well). Their exposure to this system came not through formal training, but through participation in natural social situations with human use of the system embedded in these meaningful situations, thus replicating in important ways the language-learning environment of human children. The enculturated chimpanzees were: Panpanzee, female, age 4 years, 11 months; Panbanisha, female, age 5 years; and Kanzi, male, age 10 years, 1 month. Panpanzee is a common chimpanzee (*Pan troglodytes*); Panbanisha and Kanzi are bonobos (*Pan paniscus*).

The other three chimpanzees – the mother-reared chimpanzees – were reared in a more typical captive environment, although they had much more interaction with humans than most wild or even captive chimpanzees. These chimpanzees were raised by their biological mothers (one in the wild for some years) and had very limited experience with human-like communication systems. The mother-reared chimpanzees were: Tamuli, female, age 3 years, 7 months; Mercury, male, age 4 years; and Matata, female, age 21 years, 5 months. As in the case of the enculturated chimpanzees, one of the mother-reared chimpanzees, Mercury, is of the species *Pan troglodytes* and two, Tamuli and Matata, are of the species *Pan paniscus*. The two chimpanzee groups were thus matched on the basis of gender and species and, as closely as possible, on the basis of age; that is, there was one older bonobo in each group and one of each species represented in the two younger chimpanzees in each group (with two females and one male in each group as well).

Human Ss were a randomly chosen subset of the 18-month-old Ss originally observed by Tomasello, Savage-Rumbaugh, and Kruger (1993). This group was comprised of six white, middle-class children: three males and three females (mean age = 18.4 months, range = 17.3 to 19.2 months). They were recruited by telephone from the Emory University Psychology Department Subjects file, which

© Blackwell Publishers Ltd 1995

contains the names of children whose parents voluntarily responded to a letter soliciting their future participation in child development studies. Eighteen months was chosen as the target child age because it is an age at which children begin to spend significant amounts of time in joint attentional interaction with adults (Bakeman & Adamson, 1984), but it is still before the children are too verbal, at which time visual joint attention may decline.

Materials and Observational Procedure

Ss were filmed while participating in the Tomasello, Savage-Rumbaugh, and Kruger (1993) study of imitative learning, which focused solely on Ss' abilities to reproduce observed actions on objects. Human Ss were tested in a university playroom. The experimenter (E) and child sat on the floor facing each other, with a research assistant to the side of E. The child's mother remained nearby, but was instructed not to participate. (Occasionally, however, she entered into the interaction.) The chimpanzees' sessions were conducted in various rooms or enclosures at the Language Research Center, where they sat either on the floor or on a small chair facing a familiar caretaker (the third author) who served as E. A research assistant videotaped the sessions. On occasion, other chimpanzees were also present in the room, but only rarely interfered with the experimental procedure.

Materials for the study were sixteen objects, mostly unusual toys or hardware items, that Ss had a chance to experience first in freeplay. There were two experimental sessions, spaced 48 hours apart, during which a total of 28 imitation trials were given (two actions modeled for 12 objects, one action for 4 objects; see Tomasello, Savage-Rumbaugh, & Kruger, 1993, for details). Each trial consisted of a model and response period. Model periods consisted of that time during which E actually was modeling the target action for S at the beginning of each imitation trial. Response periods began immediately after the models as E gave the object to S (or S took it) and included only the time Ss were engaged with the object and on task. This division is important for current purposes not only because Ss were in possession of the object during the response but not the model periods, but also because the character of E-S interaction differed between the two periods. During model periods, Ss were expected to pay attention to the demonstration of the target action and thus E usually was very directive, using verbal and nonverbal means to secure Ss' attention. During response periods, on the other hand, with the exception of occasional prompts and directives to do the target action, E generally watched the Ss' behavior in a less directive manner.

Coding Procedure

For purposes of the present study, the social interactions that took place during the 28 imitation trials were coded by the first author. Periods of time during which Ss were off task were not coded. The behavior of the Ss was coded in terms of their second-by-second visual attention and behavior toward both the object of focus (i.e., the object of the ongoing experimental trial) and E. Thus, for each second, Ss' visual attention was coded as *Look to the Object (LO)* or *Look to the Experimenter's Face (LE)* for at least part of the second[1], or *Neither* of the above. Ss' behavior to the object was coded according to whether Ss reproduced the *Means Only* of the target action (without success in producing the end result, e.g., turning the crank on a reel but insufficiently for retrieval of the toy attached to the string being reeled); the *End Only* (using undemonstrated means, e.g., retrieved the toy by grabbing instead of reeling); *Both Means and End* (e.g., reeled in the toy); or *Neither Means nor End* (i.e., was not engaged in attempting to reproduce the modeled action; see Tomasello, Savage-Rumbaugh, & Kruger, 1993 for details). Ss' nonverbal behavior to E was coded as *Imperative* (i.e., serving to direct E's behavior; this included requests for the object and requests for specific actions on the object), *Declarative* (i.e., serving to direct E's attention to an object by, for example, pointing, or holding up and showing), or *Neither* of the above.

At times, due to camera angles, key aspects of the interaction (e.g., the S's or E's eyes or the object) were not visible to coders. On many occasions when Ss' eyes were not clearly discernable, however, direction of gaze could nevertheless be determined reliably by head orientation. Similarly, when the object or E was temporarily off screen, Ss' looks to the object or to E's face often could be determined through knowledge of the object's or E's position during this period. Occasionally, looks and/or behavior could not be determined by the above methods. These off-screen periods were not used in any of the analyses. On average, 93% of the videotaped time was codable (96% for children, 89% for enculturated chimpanzees, and 90% for mother-reared chimpanzees). The mean time coded for the three groups was 21.6 minutes for children, 16.1 minutes for enculturated chimpanzees, and 18.9 minutes for mother-reared chimpanzees.

Following this coding, joint attentional episodes (JAEs) were identified on the basis of a formula involving Ss' alternating looks to the object and to E. JAEs were operationally defined as episodes beginning with a look to E's face (LE), given that there was engagement with an object. Engagement with an object required a look to the object (LO) at some point during the preceding or following 3 sec of the look to E. JAEs ended when the S looked away from both the object and E for 3 or more consecutive seconds.[2]

Reliability

The first author and a trained research assistant coded together until roughly 90% agreement was reached on all measures. For purposes of computing reliability estimates, one entire videotape from each of the three subject groups was selected randomly and coded independently by the first author and the research assistant (approximately 25% of the observations). Based on a second-by-second comparison, Cohen's Kappas ranged from .62 to .84 for each coding category, with a mean Kappa of .77 across all categories – in the good to excellent range as given by Bakeman & Gottman (1986). The research assistant was not blind to group membership, but was generally blind to the hypotheses of the study.

Results

Results are presented in two sections. First, Ss' joint attention (including looks to E, looks to the object, and more extended periods of joint attentional engagement) and gestural communication are examined. Second, the relationship between joint attention and imitative learning for all Ss is explored.

© Blackwell Publishers Ltd 1995

Group Differences in Joint Attention and Gestural Communication

The first set of analyses concerned various measures of joint attention and gestural communication. The small number of Ss in the groups made inferential statistics problematic. We thus present and describe a scatterplot representing the scores of individual Ss for each measure. Because the precise length of trials and sessions varied for individuals, measures involving frequencies are presented as proportions, correcting for these variations.

Looks to the Experimenter's Face (LEs). Figure 1 presents the proportion of time Ss spent looking to E during model and response periods. During the model periods Ss spent very little time looking to E's face (they mostly were looking at the objects on which actions were being demonstrated – see next section). Nevertheless, children engaged in LEs for somewhat more time than either chimpanzee group: 5 of the 6 chimpanzees spent either 2% or 3% of the model period looking to E whereas 4 of the 6 children spent from 6% to 15% of the model period looking to E. During response periods, when Ss were much more free to look where they pleased, group differences were even more striking. All 6 chimpanzees spent from 4% to 11% of the response period looking to E whereas 5 of the 6 children spent from 14% to 19% of the response period looking to E – almost no overlap in distributions. Enculturated and mother-reared chimpanzees behaved very similarly on this measure during both model and response periods.

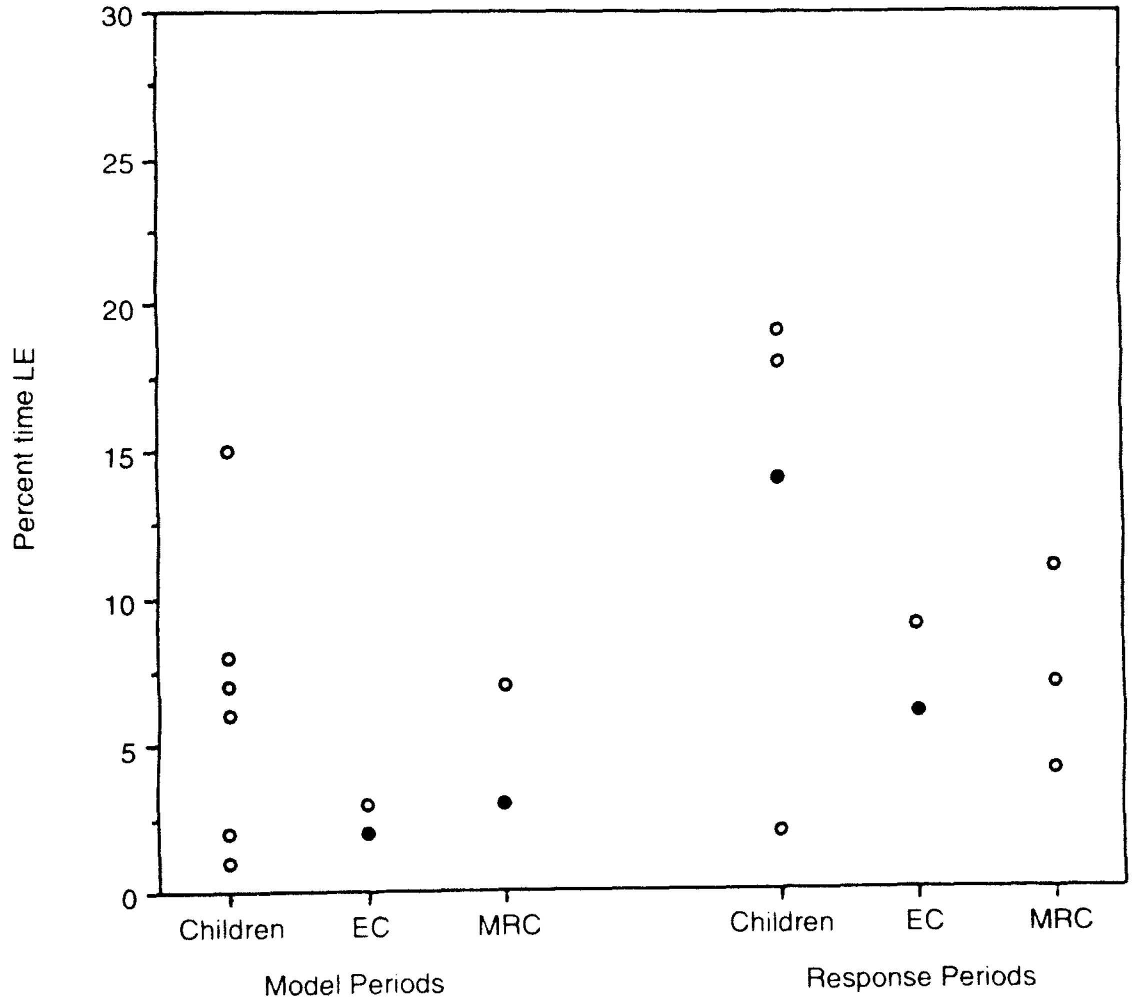

Figure 1. Proportion of time spent looking at the experimenter's face (LE) during model and response periods for children, enculturated chimpanzees (EC), and mother-reared chimpanzees (MRC).

Note: Overlapping scores are indicated by dark circles.

© Blackwell Publishers Ltd. 1995

Also different among groups was the average length of LE (see Figure 2). During model periods, 5 of the 6 children had average LEs ranging from 1.25 to 1.67 sec, whereas all 6 chimpanzees' average LEs ranged from 1.00 to 1.17 sec – again almost no overlap in distributions. During response periods, there was no overlap in distributions: all children were higher on this measure (1.71 to 2.70 sec) than any individual from the two chimpanzee groups (1.11 to 1.32 sec). In both cases, chimpanzee groups had almost identical distributions. Overall, then, what this pattern of results for LEs strongly suggests is that there is a species difference in the average length of looks to E's face between human children and chimpanzees, no matter how the chimpanzees were raised.

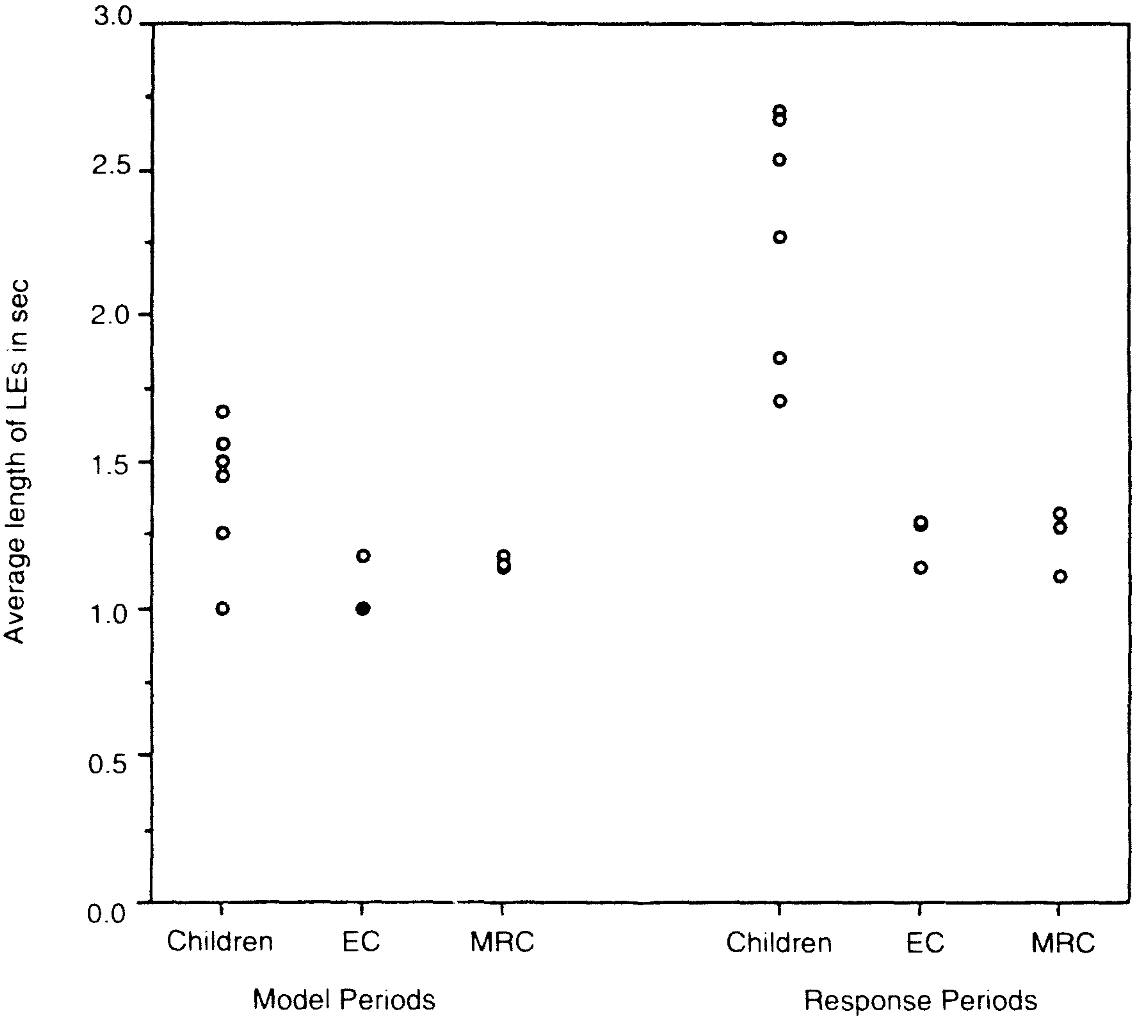

Figure 2. Average length of looks to the experimenter's face (LEs) during model and response periods for children, enculturated chimpanzees (EC), and mother-reared chimpanzees (MRC).

Note: Overlapping scores are indicated by dark circles.

Looks to the Object (LOs). Figure 3 presents the proportion of time Ss spent looking at the object during model and response periods. During the model periods, when E was trying to focus Ss' attention on the demonstration, children and enculturated chimpanzees looked at the object for a greater percentage of time than did mother-reared chimpanzees. The score of only one of the three mother-reared chimpanzees fell within the distribution of the nine children and enculturated chimpanzees (the other two falling well below). During response periods, when Ss were acting on the object, group differences in amount of time looking at the object were less clear-cut than in the model period. It is noteworthy that the children and enculturated chimpanzees looked at the object 10% more often when

© Blackwell Publishers Ltd 1995

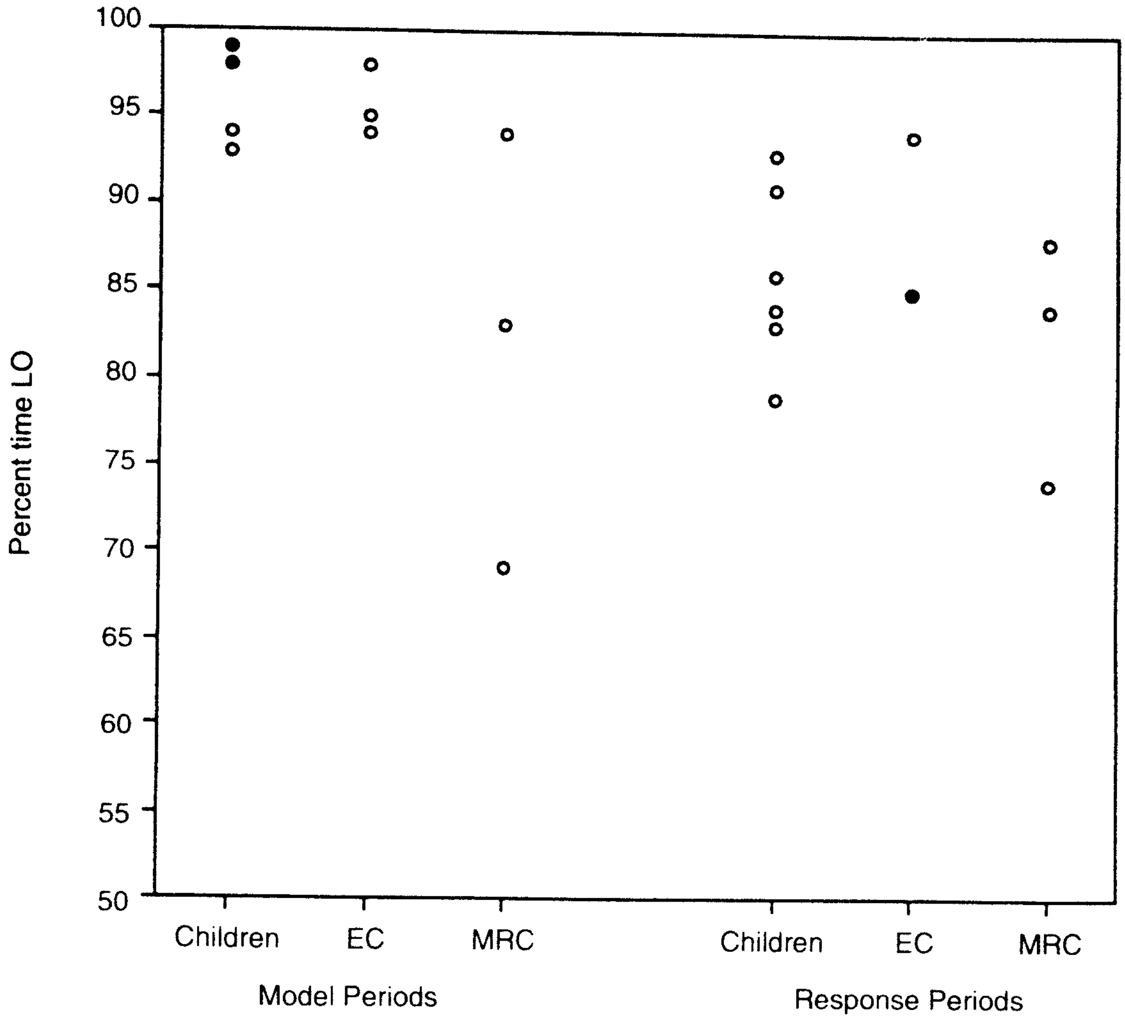

Figure 3. Proportion of time spent looking at the object (LO) during model and response periods for children, enculturated chimpanzees (EC), and mother-reared chimpanzees (MRC).

Note: Overlapping scores are indicated by dark circles.

they were encouraged to do so by E (i.e., in the model periods), whereas the mother-reared chimpanzees looked at the object equal proportions of time in both model and response periods.

Determining the average length of LOs, in a manner similar to the analyses for LEs, was not feasible in this study because Ss, especially chimpanzees, shifted their attention so often while looking at objects (and the amount of time spent looking at the object was over 80% of the codable time). We therefore chose to use a sampling methodology to determine the average length of LO for the three groups. For each S, we identified and determined the average length of the first 10 LOs beginning with the same trial (i.e., the response period of the third action modeled). Inspection of the values for the three groups (see Figure 4) shows that there is very little overlap in the groups' distributions: children's LOs ranged from 3.62 to 12.45 sec, enculturated chimpanzees' LOs ranged from 2.71 to 4.36 sec (one S overlaps with children), and mother-reared chimpanzees' LOs ranged from 1.86 to 2.50 sec (no overlap with other groups).

Overall, then, what the pattern of results for LOs suggests is that, when encouraged by humans to do so, enculturated chimpanzees attend to objects in a manner similar to human children (and dissimilar to their mother-reared conspecifics). In periods in which humans are not encouraging attention, there are still some effects of the enculturation process, especially in the tendency of chimpanzees to employ looks of longer duration to objects.

© Blackwell Publishers Ltd. 1995

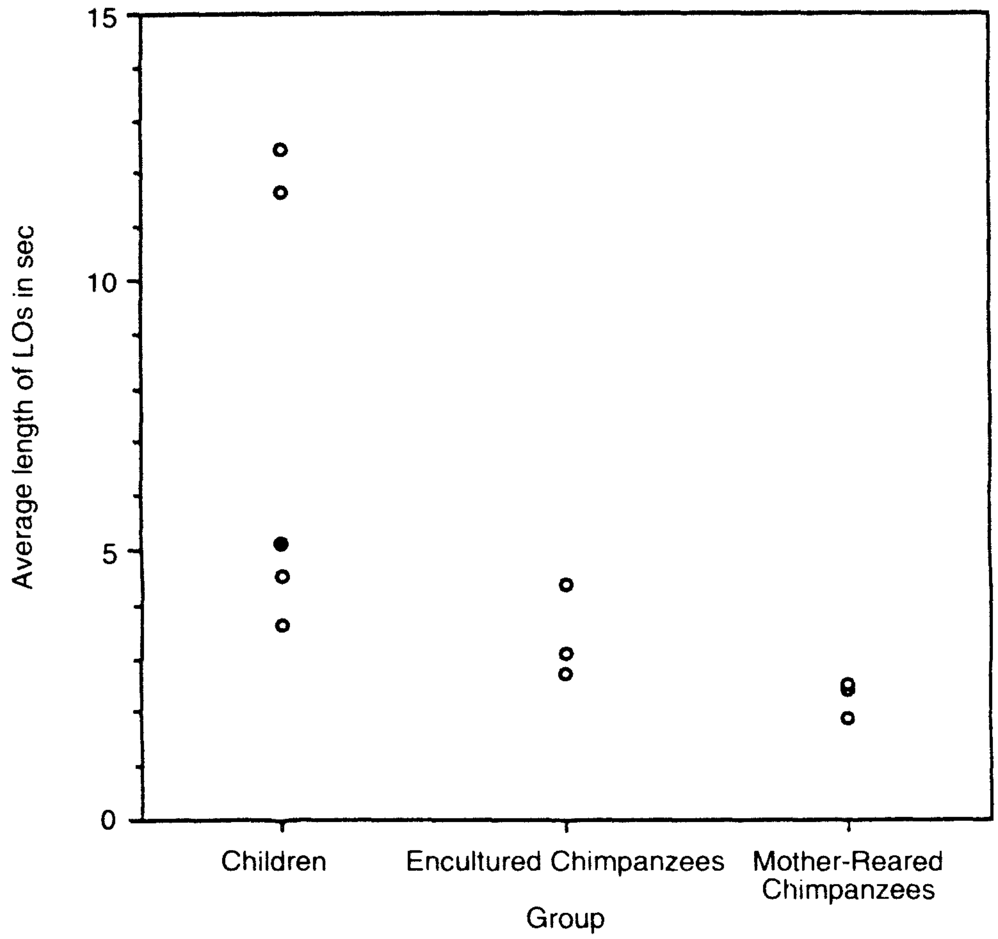

Figure 4. Average length of looks to the object (LOs) during response periods for children, enculturated chimpanzees, and mother-reared chimpanzees.

Note: Overlapping scores are indicated by dark circles.

Joint Attentional Episodes (JAEs). Figure 5 presents the proportion of time Ss spent in JAEs during the response periods. (JAEs were examined for the response periods alone because the model periods consisted only of those seconds during which E actually was performing the target action and thus often were a series of brief periods lasting a few seconds each.) Examination of Figure 5 reveals group differences in the proportion of time in JAEs across the three groups. Except for one child who engaged in very little joint attention, none of the six chimpanzees had a high enough score to overlap with the scores of the children.[3]

There also were group differences in the average length of JAEs (see Figure 6). In this case there is some indication of group differences among all three groups. Enculturated chimpanzees' average lengths of JAEs (range = 6.68 to 13.63 sec) fell in between those of children (11.26 to 22.65 sec) and mother-reared chimpanzees (6.75 to 10.42 sec).

The overall picture of group differences in JAEs is thus a relatively consistent ordering: children spend a great deal of time engaging in JAEs and engage in relatively long JAEs, mother-reared chimpanzees spend much less time engaging in JAEs and engage in very short JAEs, and enculturated chimpanzees fall in between children and mother-reared chimpanzees on both measures.

Control Analyses. As reported in the method section, there were some differences in the way the children and chimpanzees were videotaped: chimpanzees' eyes and/or heads were off-screen slightly more often than were children's (10% to 4%

© Blackwell Publishers Ltd 1995

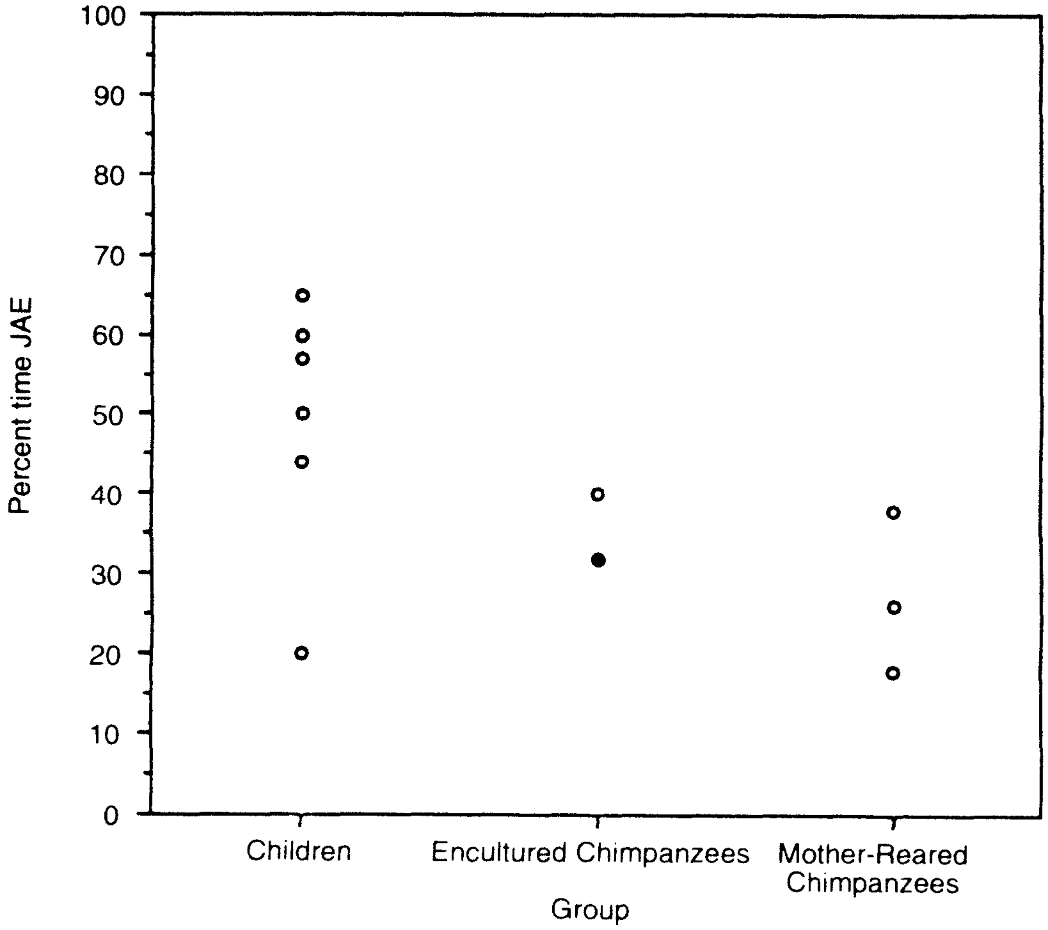

Figure 5. Proportion of time spent engaging in joint attentional episodes (JAEs) (including JAEs with children's mothers) during response periods.

Note: Overlapping scores are indicated by dark circles.

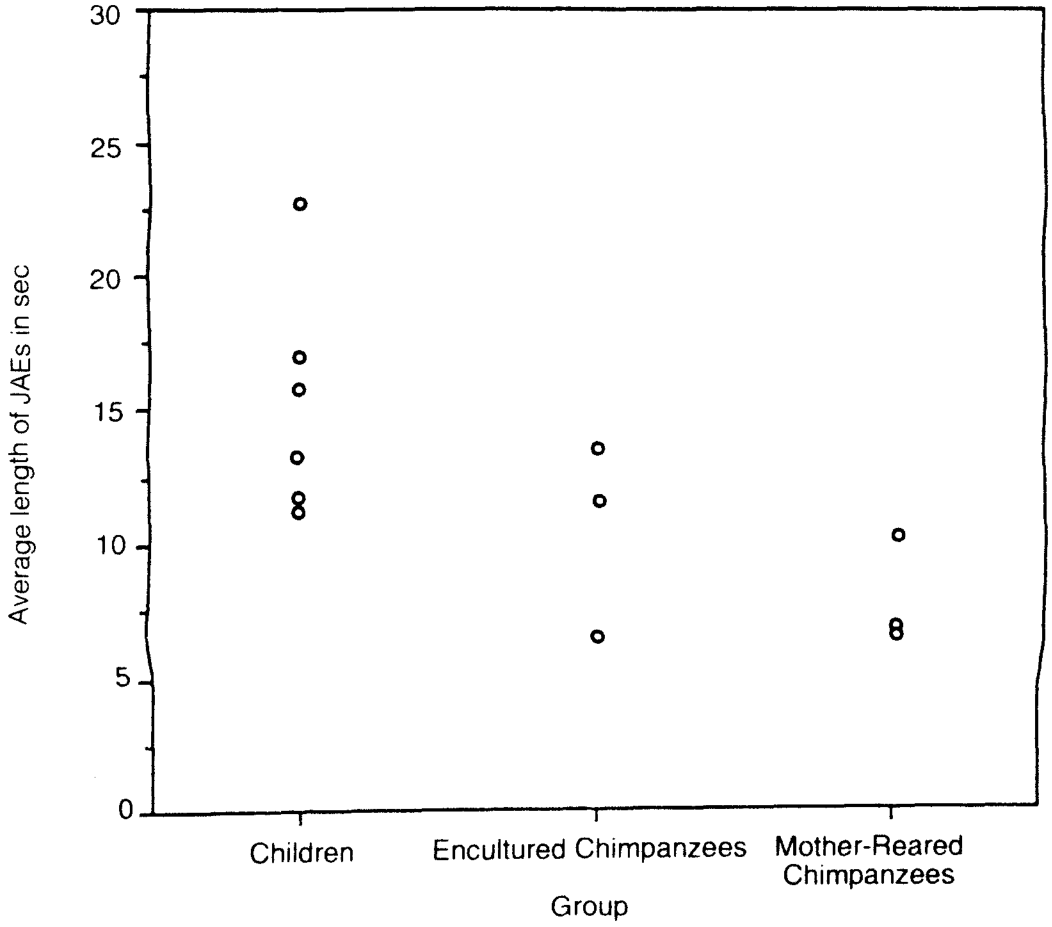

Figure 6. Average length of joint attentional episodes (JAEs) (including JAEs with children's mothers) during response periods.

© Blackwell Publishers Ltd. 1995

of the time, respectively). This raises the possibility of an artifactual explanation for the results concerning the average length of LEs and JAEs. That is, it may have been the case that chimpanzees' LEs and JAEs were shorter because they were interrupted more often by uncodable periods. To test this possibility, the average lengths of LEs and JAEs were recalculated excluding all interrupted LEs and JAEs for all Ss. Results were nearly identical to those reported above. (In fact, the exclusion of interrupted JAEs slightly increased the difference in average length between children's and chimpanzees' JAEs.) Group differences in average length of LEs and JAEs thus cannot be attributed to group differences in videotaping procedures.

Gestural Communication. Ss' attempts to direct E's behavior (imperatives) or attention (declaratives) by means of gestural communication also were of interest. Unfortunately, the current experimental situation was not ideal for assessing Ss' intentional communication as Ss were under the direction of E for much of the time as she tried to get them to watch the model or to perform the modeled action. In addition, objects were first introduced to Ss during the freeplay periods preceding the model and response periods analyzed above. Objects were thus newest then so that period may have provided the clearest opportunities for showing, pointing, and so forth. For this reason we decided to supplement the numbers from the model and response periods with the intentional communication that occurred in the freeplay periods.[4] Whereas children, enculturated chimpanzees, and mother-reared chimpanzees produced similar numbers of imperatives, there were striking group differences with regard to declaratives (see Figure 7). Children as a group produced 72 declaratives (range for individual children = 7 to 23)[5] whereas mother-reared chimpanzees produced no declaratives at all. Enculturated chimpanzees as a group produced 2 gestures that were interpreted as declaratives: one was a point to part of an object (although the chimpanzee's intentions were not totally clear) and one was a gesture by another enculturated chimpanzee resembling a show (i.e., the chimpanzee held up an object draped around his ankle in the direction of E's face).

Gender, Age, and Chimpanzee Species Differences. There were no marked gender differences in any of the groups on any of the above measures. Likewise, there were no species differences between *Pan troglodytes* and *Pan paniscus* on any measure. There was one minor age difference for the chimpanzees, however. The two older (10–21.5 years) chimpanzees engaged in slightly shorter JAEs than the four younger (3.5–5 years) chimpanzees.

Joint Attention and Imitative Learning

The second set of analyses investigated the relationship between Ss' joint attentional and imitative learning abilities. Measures of joint attention in this case included LOs and LEs from both model and response periods and JAEs from response periods only (as in the group comparisons). Measures of imitative learning consisted of the proportion of trials in which Ss reproduced the Means Only, End Only, Neither Means nor End, and Both Means and End of E's demonstration. In addition, a single score across trials (called total imitation) was calculated for each S by assigning a score of 0 for Neither trials, 1 for Means Only and End Only trials, and 2 for Both Means and End trials, and then averaging these scores

© Blackwell Publishers Ltd 1995

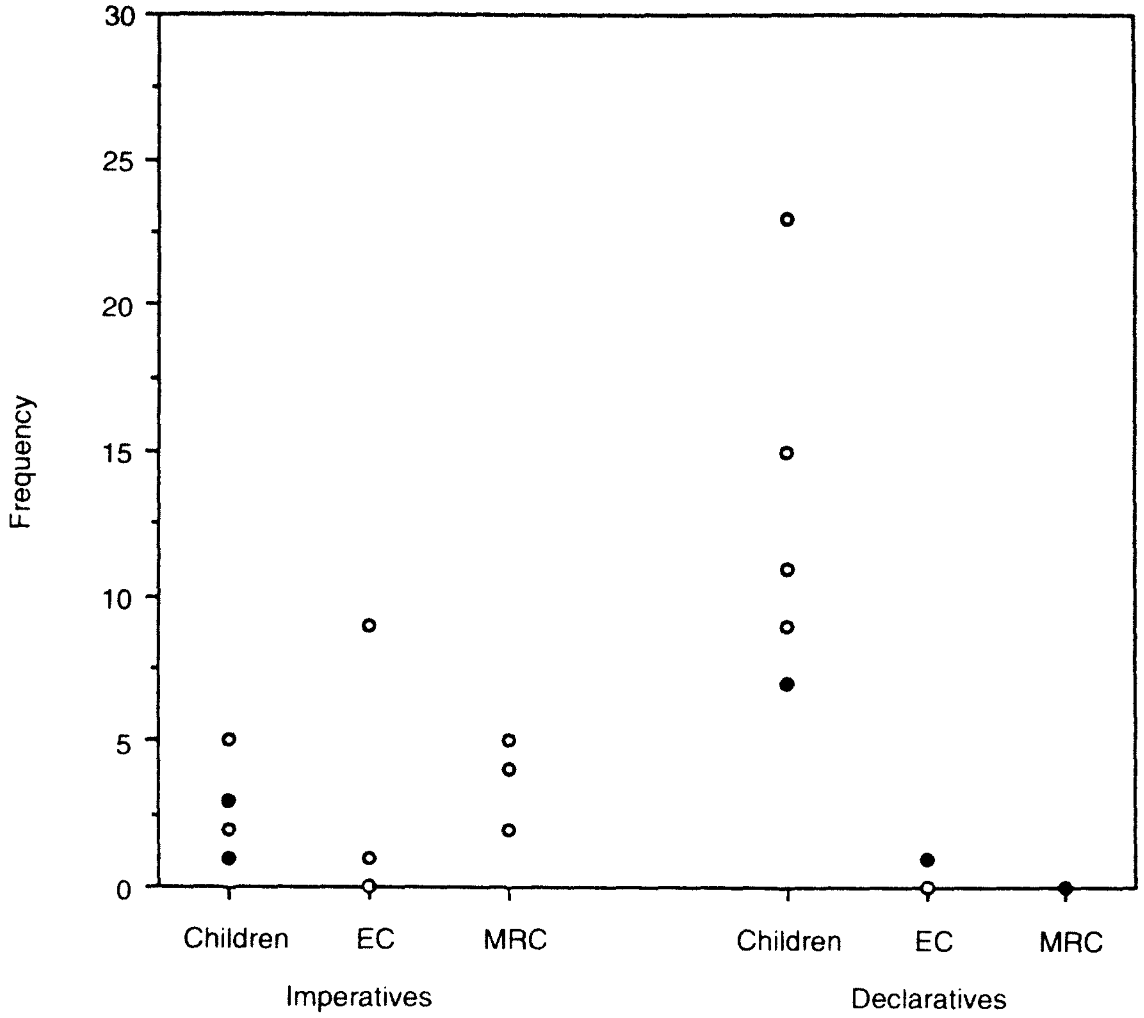

Figure 7. Frequency of imperative and declarative gestures during model periods, response periods, and freeplay for children, enculturated chimpanzees (EC), and mother-reared chimpanzees (MRC).

Note: Overlapping scores are indicated by dark circles.

(excluding trials for which the S performed the target action in the freeplay period as these precluded the possibility of imitative learning). In this case inferential statistics, in the form of Pearson Product Moment correlations, were deemed appropriate as there was a total of 12 Ss entering into each correlation.

Model Periods. Table 1 presents the correlations between Ss' imitative learning scores and their LOs and LEs during the model periods. The most striking result involves Ss' LOs during the model periods and their subsequent attempts at imitation. The greater proportion of time Ss looked at the object during model periods, the better they did imitatively. There were significant positive correlations between the percentage of time spent looking at the object and both total imitation score, $r = .63$, and the percentage of trials in which Ss successfully reproduced Both the Means and End result of the target action, $r = .55$. There also was a negative correlation between percent time looking at the object and the percentage of trials in which Ss reproduced Neither the Means nor the End result of the target action, $r = -.50$ ($p < .05$ in all cases). These relationships indicate a positive association between paying attention to the modeled action and the ability to later reproduce it.

An interesting pattern of results emerged in the case of LEs. As can be seen in Table 1, there was a significant positive correlation between average length of LE during models and the percentage of trials in which the S reproduced the Means

Table 1. Correlation Matrix: Looks to the Object (LOs) and Looks to the Experimenter (LEs) During Model Periods with Imitation Scores

	Joint attention		
Imitation	Percent time LO	Percent time LE	Average length LE
Percent of trials: Neither Means nor End	–.50*	.04	–.13
Percent of trials: Means Only	.30	.36	.67*
Percent of trials: End Only	.06	–.40	–.48
Percent of trials: Both Means and End	.55*	–.13	–.05
Total imitation score	.63*	.00	.13

*$p < .05$

Only of the target action. Conversely, there was a negative correlation between average length of LEs during models and the percentage of trials in which the S reproduced the End Only of the target action (this same pattern held nonsignificantly when proportion of total time spent looking at E was used as a measure). Thus, the longer Ss' LEs (and, to a lesser extent, the more time Ss spent looking to E's face) during models, the more likely Ss were to reproduce the means only of the target action and the less likely they were to reproduce the end only of the target action. This result makes sense in the current context because the means of the target action is defined in terms of E's particular behaviors, which Ss would have to look at E to discern (as E's face perhaps would provide additional information about what she was trying to do), whereas the end result of the target action is defined in terms of a change of state of the object (which could be seen without looking at E's behavior or face at all).

Response Periods. Table 2 presents these same correlations between Ss' imitative learning scores and their LOs and LEs, with two measures of JAEs in addition, during the response periods. Unlike for the model periods, LOs during response periods were not significantly correlated with imitative learning. This indicates quite simply that looking at the object while attempting to reproduce the modeled action was not particularly helpful. Consistent with findings from the model periods, however, both LE measures were positively correlated with the percentage of trials in which the S reproduced the Means Only of the modeled action, and negatively correlated with the percentage of trials in which the S reproduced the End Only of the modeled action. The meaning of these correlations may be slightly different from that of the correlations from the model periods, however. LEs during the response periods came after the model was over, while the S was attempting to reproduce the modeled action. (And this is what Means Only responses mostly were: unsuccessful attempts at reproducing the modeled action.) The positive correlation between LEs and Means Only may thus indicate that looking to E for encouragement or feedback during response periods occurred most often

© Blackwell Publishers Ltd 1995

Table 2. Correlation Matrix: Looks to the Object (LOs), Looks to the Experimenter (LEs), and Joint Attention (JAEs) during Response Periods with Imitation Scores

	Joint attention				
Imitation	Percent time LO	Percent time LE	Average length LE	Percent time JAE	Average length JAE
Percent of trials: Neither Means nor End	–.24	–.05	–.34	–.28	–.51*
Percent of trials: Means Only	.06	.49	.61*	.47	.51*
Percent of trials: End Only	.20	–.51*	–.44	–.35	–.09
Percent of trials: Both Means and End	.24	.00	.31	.28	.46
Total imitation score	.27	.15	.41	.37	.55*

* $p < .05$

during trials in which Ss were attempting unsuccessfully to reproduce the modeled action. Conversely, End Only trials were those in which Ss knew how to change the state of the object and did so with no need of guidance or encouragement from E – and thus the negative correlation with LEs.

What was most strongly associated with overall imitative performance during the response periods was the average length of JAEs. As can be seen in Table 2, average length of JAEs during response periods was positively correlated with total imitation score, $r = .55$, $p < .05$ (and nonsignificantly with the proportion of Both Means and End trials). It also was negatively correlated with the percentage of trials in which the S reproduced Neither the Means nor End of the modeled action, $r = -.51$, $p < .05$. These relationships indicate that during the response period, when the model was over and the S possessed the object and was attempting to reproduce the modeled action, the alternation or coordination of attention to both object and E was associated with higher levels of imitative learning.

Group Differences in the Relationship Between Joint Attention and Imitative Learning. The correlational analyses just presented included all 12 Ss across all 3 groups. They thus do not address possible group differences in the relationship between Ss' joint attentional and imitative learning abilities. Examination of the scatterplots for the two strongest correlations of the overall analyses revealed the following. For the response periods, the positive correlation between the average length of JAEs and total imitation score held not only for the Ss as a whole, but also within each of the groups individually. For the model periods, however, the overall correlation between LOs and total imitation score was mainly a function of group differences: mother-reared chimpanzees were substantially lower than the other two groups on both of these measures, and all nine of the child and enculturated chimpanzee Ss were bunched quite tightly together with relatively high values on both (with no correlation within groups). It would thus seem that large

© Blackwell Publishers Ltd. 1995

variations in attention to the object during models are associated with strong imitative performance, but small variations do not matter.

Also of interest for current concerns was the relation of Ss' LEs during the model periods with their imitative performance. This was of theoretical interest and importance because it might be argued that in order to discern what E was trying to do with the target object and how she was trying to do it – and thus to be able to reproduce both end and means – Ss would need to look to both the object and to E during the model period. As can be seen in Table 1, however, no significant correlation was found between LEs and total imitation for the Ss as a whole during the model periods; the positive correlation was with LOs only. We suspected, however, that a linear model may not be the best model to apply in this case. Ss who looked to E too much during the model would be diminishing their time looking to the object, and time spent looking to the object was positively correlated with total imitation. Therefore, in recognition of this complementarity between LOs and LEs in the model periods, we took the following analytic approach. We first distinguished those trials for which the S looked to E's face at least once during the model period from those trials for which there were no looks to E's face during the model period. We then marked each trial for its outcome in imitative learning: Means Only, End Only, Neither, or Both. Each S's score was then computed by taking the trials in which there was at least one LE during the model period and determining the proportion of those trials that resulted in each type of imitative learning. The same procedure was then followed for trials in which there was not an LE during the model period. The mean proportions for groups were then calculated from these individual scores, and then compared for each type of imitative learning outcome separately.

Table 3 presents the results of this analysis. As can be seen by comparing Tables 3a and 3b, looking to E's face or not during the model period made very little difference in how well children and mother-reared chimpanzees performed on the imitation task. On the other hand, looking to E during the model period did make a difference to enculturated chimpanzees. When they looked to E's face at least once during the model period, enculturated chimpanzees succeeded in reproducing Both Means and End of the modeled action much more often (60%) than when they did not look (37%). Similarly, enculturated chimpanzees reproduced Neither Means nor End of the modeled action more often when they did not look to E's face during the model period (35%) than when they did (17%). It would thus seem that looking to the face of E at least once during the model period is important to the reproduction of both end and means of the target action for enculturated chimpanzees in a way that it is not for the other two groups – perhaps because their imitative learning skills were of a slightly different nature than those of the other two groups.

Discussion

The groups we observed in the current study differed from one another in the following ways. The most pronounced species differences were: (1) the vast majority of looks to E's face by chimpanzees of both groups were brief glances, whereas the majority of children's looks to E were more long-lasting; and (2) children used declarative gestures much more frequently than chimpanzees. The clearest effect of the enculturation process was in the way chimpanzees attended to objects and

© Blackwell Publishers Ltd 1995

Table 3a. Proportion of Trials Resulting in Each Type of Imitative Learning when Subject Looked to the Experimenter's Face at Least Once during the Model

	Children	Enculturated chimpanzees	Mother–reared chimpanzees
Neither means nor end	44.8	16.7	82.3
Means only	26.5	16.7	11.0
End only	2.7	6.7	3.3
Both means and end	25.5	60.0	3.3

Table 3b. Proportion of Trials Resulting in Each Type of Imitative Learning when Subject did not Look to the Experimenter's Face during the Model

	Children	Enculturated chimpanzees	Mother–reared chimpanzees
Neither means nor end	40.2	35.3	76.0
Means only	27.3	11.7	11.7
End only	4.7	16.3	4.3
Both means and end	27.8	36.7	8.0

interacted with humans around objects: (1) enculturated chimpanzees looked at the object for more time than mother-reared chimpanzees, both when E was demonstrating actions on objects and encouraging Ss to watch and when they were left to their own devices in the response period; and (2) enculturated chimpanzees engaged with E and objects in periods of joint attentional focus for longer periods than mother-reared chimpanzees.

The explanation for these results lies in the nature of the enculturation process. A major goal of the humans who interact with Kanzi, Panbanisha, and Panpanzee on a daily basis is to encourage them in play with objects, to teach them new actions to perform on objects, and, in general, to engage them with the material culture that surrounds them. The current results would suggest that chimpanzees are sensitive to this type of social direction. On the other hand, looks to the faces of humans and gestures to humans are not socialized in this same way. Thus, while human caretakers do encourage Kanzi *et al.* to look to them, the average length of those looks is not considered important; a brief glance usually is sufficient to convince the human of the chimpanzee's attention. In human infants' long history of face-to-face interactions with adults, longer looks to the face are explicitly encouraged through smiles, attention-getting devices, and other forms of 'protoconversations' (Trevarthen, 1979). The overall point is that it is how chimpanzees attend to objects, not specifically to people through looks or gestures, that is the major focus of the enculturation process, and thus it is there that we see the biggest differences between the enculturated and mother-reared chimpanzees.

© Blackwell Publishers Ltd. 1995

Interpretation of these results should be tempered by three considerations. First, an important caveat to the findings is that the child and chimpanzee Ss interacted with different Es. Because many of the measures of direct concern were interactive in nature, this difference could potentially have had a significant effect on our group comparisons. There is really no solution to this dilemma, since we will always need to have human Es with human children, and thus E will always differ in species from the chimpanzees; also, a caretaker familiar to the chimpanzees will very likely be unfamiliar to children. It is important to point out, however, that we are in general placing more importance on the differences we have found between the two chimpanzee groups than we are on species differences, and the two chimpanzee groups interacted with the same E. We should also note that our findings may be dependent to some degree on the nature of the specific task Ss were engaged in – they were participating in a task in which E had things she wanted them to do – and that in a different, less directive setting the nature of the joint attentional interactions for all groups of Ss might be different.

A second important consideration is that in the investigation of joint attention, it is important to know something of the Ss' intentions when they look to the object and the partner. Because joint attention is defined in terms of a *sharing* of attention (Bakeman & Adamson, 1984; Tomasello, 1995), the look to E should be a look aimed at engaging E or assessing E's attitude towards the *object*. If, for example, an individual is playing with a toy and looks to E to see if E is preparing to reward or punish her, that is not a sharing situation because the individual is not monitoring E's attention to the *object*. Along these same lines, if an individual looks up from a toy to investigate a noise E has made and then goes back to the toy, that is best called something like 'alternating attention', since again this is not an interaction in which both participants are monitoring the other's attention to an object. The important point is that because of the 'objective' manner of coding in the current study – only direction of gaze was coded – it is likely that Ss in all groups had some interactions coded as joint attention that were not reflective of a true coordination of attention. The large number of very brief looks to E by the chimpanzees might seem to indicate that many of their looks were simply 'checking' on the adult or alternating attention, although it is also possible that the generally quicker attention deployment of apes allowed them to monitor E's attentional state with brief glances.

Finally, we should point out that the mother-reared chimpanzees in this study had had much human contact. There is evidence that by 30 days of life, nursery-reared chimpanzees already perform at 'higher and more human-like levels' than mother-reared chimpanzees on several measures of visual orientation (Bard, Platzman, Lester, & Suomi, 1992). Whereas none of the mother-reared chimpanzees in this study was raised in a nursery or given the extensive human contact or rich environment of the enculturated chimpanzees, they nonetheless were exposed quite frequently to human caretakers, in some cases from a fairly early age (Savage-Rumbaugh *et al.*, 1983; Savage-Rumbaugh, 1993). The mother-reared chimpanzees in this study therefore are very likely not representative of chimpanzees raised in their natural habitats, or even in more typical captive situations, and it is thus possible that if the enculturated chimpanzees in this study were compared with chimpanzees more typical of their species than our mother-reared group, even larger group differences would have been observed.

We thus believe that, overall, our findings of group differences (including those

© Blackwell Publishers Ltd 1995

reported in Tomasello, Savage-Rumbaugh, & Kruger, 1993) reflect something very important about the role of the social-cultural environment in the ontogeny of human-like skills of joint attention and imitative learning. At least some species differences between chimpanzees and human children in these skills are due to the fact that humans are raised in rich cultural environments. Specifically, we believe that such things as intentional instruction from adults, interaction with objects in the presence of others who encourage attention to those objects, reinforcement for imitating adult actions, communication with adults using conventional symbols, and being treated as an intentional agent oneself are all important ingredients in coming to understand other persons as intentional agents. This understanding is foundational in the ontogeny of joint attentional, imitative, and communicative skills, for both chimpanzees and humans (Tomasello, Kruger, & Ratner, 1993; Tomasello, 1995).

It is also important that in the current study we found a positive relation between joint attention and imitative learning. If imitative learning involves reproducing not only the end result of an action on an object but also its behavioral means, it would seem reasonable that the imitator ought to pay attention to both the object whose state is changing and the person whose behavior is changing it. Consistent with this view, we found that across all Ss the percentage of time spent looking at the object during model periods was positively correlated with total imitation score and that looking at least once to E's face during the model period was positively related to the reproduction of both end and means of the target action for the enculturated chimpanzees. We hypothesize that this is because information about E's intentional state as expressed in her face (e.g., where she is looking, whether she seems to be happy or frustrated with the result of her actions, etc.) is useful for reproducing her actions in the way she produced them if one's imitative learning skills are of a certain type. The fact that chimpanzees only show sophisticated imitative learning skills when raised in human-like cultural environments might suggest that they need to work especially hard in such tasks and that lapses of attention would lead to failure more often than for human children. It would thus be interesting to investigate the relationship between joint attention and imitative learning in human infants at around 9 months of age, when their imitative skills first emerge, to see if the same vigilance is required for these infants.

In all, we believe that comparative studies have a crucial role to play in our understanding of both the phylogeny and ontogeny of cognitive skills. We think they are especially important in the case of social-cognitive skills, as it is rapidly becoming evident that processes of social cognition are what most clearly distinguish the cognition of primates from that of other mammalian species (see, e.g., Byrne & Whiten, 1988). Cross-fostering experiments provide crucial additional information about the role played by specific aspects of the physical and social environment in the ontogeny of primate social cognition. While such experiments are obviously not possible in the human case, their employment with apes is one of the most important sources of information available on the role of environmental variations on significant cognitive outcomes. In the current study we have attempted to use both the comparative and cross-fostering approach in order to identify not only species differences but also differences in the social-cognitive development of chimpanzees that may be attributed to differences in the role of human culture in their rearing environments.

© Blackwell Publishers Ltd. 1995

References

Bakeman, R., & Adamson, L. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interactions. *Child Development, 55,* 1278–1289.

Bakeman, R., & Gottman, J. (1986). *Observing interaction: An introduction to sequential analysis.* Cambridge: Cambridge University Press.

Bard, K. A., Platzman, K. A., Lester, B. M., & Suomi, S. J. (1992). Orientation to social and nonsocial stimuli in neonatal chimpanzees and humans. *Infant Behavior and Development, 15,* 43–56.

Bard, K. A., & Vauclair, J. (1984). The communicative context of object manipulation in ape and human adult-infant pairs. *Journal of Human Evolution, 13,* 181–190.

Bates, E. (1976). *Language and context.* New York: Academic Press.

Boesch, C. (1991). Teaching among wild chimpanzees. *Animal Behavior,* 41, 530–532.

Byrne, R., & Whiten, A. (Eds.). (1988) *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans.* Oxford: Clarendon.

Gómez, J. C. (1991). Visual behaviour as a window for reading the mind of others in primates. In A. Whiten (Ed.), *Natural theories of mind: Evolution, development and simulation of everyday mindreading* (pp. 195–207). Cambridge, MA: Basil Blackwell.

Gómez, J. C., Sarría, E., & Tamarit, J. (1993). The comparative study of early communication and theories of mind. In S. Baron-Cohen, H. Tager-Flusberg, & D. Cohen (Eds.), *Understanding other minds.* Oxford: Oxford University Press.

Hayes, K., & Hayes, C. (1952). Imitation in a home-raised chimpanzee. *Journal of Comparative and Physiological Psychology, 45,* 450–59.

Meltzoff, A. (1988a). Infant imitation and memory: Nine-months olds in immediate and deferred tests. *Child Development, 59,* 217–25.

Meltzoff, A. (1988b). Infant imitation after a one week delay: Long term memory for novel acts and multiple stimuli. *Developmental Psychology, 24,* 470–76.

Plooij, F. (1978). Some basic traits of language in wild chimpanzees? In A. Lock (Ed.), *Action, gesture and symbol: The emergence of language* (pp. 111–131). New York: Academic Press.

Savage-Rumbaugh, S. (1990). Language as a cause-effect communication system. *Philosophical Psychology, 3,* 55–76.

Savage-Rumbaugh, E. S. (1993, April). Language and animal mind. Presentation at a conference The Mental Lives of Animals. Atlanta, GA.

Savage-Rumbaugh, E. S., McDonald, K., Sevcik, R. A., Hopkins, W. D., & Rubert, E. (1986). Spontaneous symbol acquisition and communicative use by pygmy chimpanzees (*Pan paniscus*), *Journal of Experimental Psychology: General, 115* (3), 211–235.

Savage-Rumbaugh, E. S., Pate, J. L., Lawson, J., Smith, S. T., & Rosenbaum, S. (1983). Can a chimpanzee make a statement? *Journal of Experimental Psychology: General, 112* (4), 457–492.

Tomasello, M. (1990). Cultural transmission in the tool use and communicatory signaling of chimpanzees? In S. T. Parker & K. R. Gibson (Eds.), *'Language' and intelligence in monkeys and apes: Comparative developmental perspectives.* (pp. 274–309). New York: Cambridge University.

Tomasello, M. (1995). Joint attention as social learning. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origins and role in development.* Hillsdale, N.J.: Erlbaum.

Tomasello, M. (1994). The question of chimpanzee culture. In R. Wrangham, W. McGrew, F. de Waal, & P. Heltne (Eds.), *Chimpanzee cultures.* Cambridge: Harvard University Press.

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences, 16,* 495–552.

Tomasello, M., Savage-Rumbaugh, S., & Kruger, A. C. (1993). Imitative learning of actions on objects by children, chimpanzees, and enculturated chimpanzees. *Child Development, 64,* 1688–1705.

Trevarthen, C. (1979). Instincts for human understanding and for cultural cooperation. In M. von Cranach, K. Foppa, W. Lepenies, & D. Ploog (Eds.), *Human ethology: Claims and limits of a new discipline* (pp. 530–571). Cambridge: Cambridge University Press.

© Blackwell Publishers Ltd 1995

Notes

1. If Ss looked to both the object and E during the same second, an LO and an LE were recorded for that second.

2. To test the validity of our operational definition of JAEs, JAEs were also recoded using more stringent criteria: in order for a JAE to begin, the LE had to occur within 1 sec (instead of 3 sec) of the LO; in addition, episodes were terminated in the usual way and also if 10 sec elapsed without an LE. The pattern of results remained unchanged.

3. Note that this comparison uses children's JAEs with both E and their mothers. We used this measure because some children on some occasions preferred to interact with their mothers; chimpanzees had only one possible interactant and, insofar as this is possible for a human, she was a mother-like figure to them. If children's JAEs with E only are counted, the overall pattern is identical, although the group differences are not quite as large. For all subsequent analyses of JAEs, children's mothers are included, and in no case does their exclusion change the basic pattern of results.

4. When combined, the mean total time of observation for each of the three groups across model, response, and freeplay periods was very similar – within 1.6 minutes of 32 minutes.

5. As in the case of JAEs, some children preferred to show and point to things for their mothers, whereas chimpanzees had only their human caregiver present. The numbers reported reflect children's gestures to both E and their mothers. If mothers are excluded, the same pattern holds but children's total number of gestures decreases by just over one half.

© Blackwell Publishers Ltd. 1995

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.