CULTURAL LEARNING AND CREATION

Tanya Behne, Malinda Carpenter, Maria Gräfenhain, Kristin Liebal, Ulf Liszkowski, Henrike Moll, Hannes Rakoczy, Michael Tomasello, Felix Warneken, and Emily Wyman

Human children become cultural beings by learning to participate in the cultural activities and practices going on around them. Household pets grow up in the midst of these same cultural activities and practices, but they do not learn to participate in them in anything like the same way as human children. Even chimpanzees and bonobos raised in human homes and treated like human children still retain, for the most part, their species-typical social and cognitive skills without turning into cultural beings of the human kind. This difference suggests that humans are biologically adapted, in ways that other animal species are not, for becoming cultural beings by tuning in to what others around them are doing, and thereby learning from them. Moreover, on occasion, young children even create with others small-scale cultural activities and routines involving one or another form of collaboration, or even collaborative pretense. Such cultural creation would also seem to be unique to human beings, and of course cultural creation leads to ever new cultural environments in which human cognitive ontogeny takes place.

We may therefore identify two sets of human cultural skills responsible, as they work over historical and ontogenetic time, for humans' unique form of social organization: cultural learning and cultural creation. These enable humans, and only humans, to have cultures which accumulate complexities in both social practices and cognitive artifacts - creating ever new cultural niches within which developing children become mature cultural beings.

Tomasello (1999) proposed that underlying these cultural abilities was a uniquely human social-cognitive skill for understanding others as intentional agents who, like the self, attend to things and pursue goals in the environment.

The collective aspect of cultural evolution in this theory was, in an important sense, taken for granted. Uniquely human types of social engagements such as joint attention, collaborative co-operation, and symbolic communication were seen as simply emanating naturally from the understanding of others as intentional agents like the self. However, recent
research with non-human primates suggests that the origins of these skills should not be taken for granted. That is, a number of different studies have found that great apes do understand important aspects of intentional action and perception. For example, chimpanzees recognize the difference between intended actions and accidental actions and they know what others can and cannot see, in the sense of the contents of their perception (see below for more details). But they do not engage in shared activities or in processes of cultural learning and creation the way that humans do from an early age on.

There may still be some differences in the ways that human children and great apes understand intentional action and perception, especially in the understanding of the choices an actor makes in creating action plans and focusing attention (taking perspectives) on things (see below). Nevertheless, these new data have driven us to acknowledge what seems to be a clear fact: human infants and young children have special motivations and skills for engaging in shared activities that go beyond the understanding of others as intentional agents like the self. Infants look to others when interesting things happen, often just to share attention to them. Infants are motivated simply to point to things so that others will share interest in them. Infants are motivated to form shared goals and shared intentions with others in their joint activities with objects. Young children create with others pretend realities that exist only in their shared intentionality. These skills and motivations for sharing do not seem to be present in nonhuman primates to nearly the same degree, if at all (Tomasello & Carpenter, 2005). The important point in the current context is that humans' unique skills and motivations for sharing psychological states with others play a crucially important role for cultural learning and cultural creation, as they lead infants and young children to tune into and learn from what others are doing, and to create with others the kinds of novel interactive routines and cultural practices that characterize specific families and cultures - which serve to create the cultural niche within which the next generation develops.

Tomasello, Carpenter, Call, Behne, and Moll (2005) therefore proposed that uniquely human cultural skills depend on two ontogenetic pathways, each reflecting a distinct biological adaptation. The first pathway concerns the understanding of intentional action and
perception. Much of this pathway is shared among all primates, or at least among the great apes, although humans may go somewhat further along this path. The second pathway is a uniquely human motivation to share psychological states with others. This can be seen in infants' earliest social interactions, known as protoconversations, in which they exchange emotions with adults dyadically, seemingly only for the purpose of sharing. When infants begin to understand intentional action and perception as directed at outside goals and targets, near the end of the first year of life, these two strands of ontogeny come together in a new way of special relevance for cultural learning and creation. Specifically, the motivation to share now manifests itself as the sharing of intentions and attention triadically in acts of joint attention, collaborative co-operation, and symbolic communication.

In this chapter we review recent research from our laboratory relevant to this theoretical account. Firstly, we review our research on infants' and young children's understanding of intentional action and perception. Secondly we look at infants’ and young children’s ability to engage in shared activities, looking specifically at skills of joint attention, communication, collaboration and pretense. The attempt is to elucidate the ontogenetic roots of the human forms of social interaction and communication that enable uniquely human forms of cultural learning and creation. (Thus, in each section we also compare the research findings to what is known about the respective abilities of non-human apes, especially chimpanzees.)

1. Understanding intentional action and perception

1.1 Understanding goals and intentions

At some point in infancy young children come to perceive the bodily motions of other people as intentional actions. Determining precisely when this developmental transition occurs is of theoretical interest because it marks an important step in the development of their cultural skills. In this section, we review our recent research on the development of infants’ understanding of others’ intentional action, and we discuss how infants use this understanding in cultural learning. We also briefly review our research on understanding of
others’ intentional action in apes and show how their perhaps more limited understanding can explain differences in which aspects of others’ behavior they reproduce in cultural learning situations.

To be clear about what we mean by intentional action, in our view intentional action involves crucially: 1) a goal or mental representation of a desired end state, 2) an intention or plan of action the actor chooses and commits himself to in pursuit of that goal, and 3) the ability to perceive and monitor one’s actions and the environment in order to know when the state of the environment matches the goal (Tomasello et al., 2005). We will focus on the first two of these components here (see section 1.2 below for research on the third component).

**Development of infants’ understanding of others’ goals and intentions**

When studying nonverbal infants’ understanding of others’ goals, the methodological challenge is to find a way to separate the actor’s goal from the result that he or she achieves. The clearest way to do this is to study infants’ reactions to actions that are not immediately successful, because in this case infants must infer the actor’s goal even though it is not achieved (and therefore not observed). The two main categories of unsuccessful actions are failed attempts (trying) and accidents. Previously, the only studies investigating infants’ understanding of others’ failed attempts and accidents used imitation as a response measure. These studies found that 14- and 15-month-old and older infants showed an understanding of successful versus unsuccessful and intentional versus accidental actions, that is, that they could determine whether the actor’s goal matched the external result produced and responded appropriately (Carpenter, Akhtar, & Tomasello, 1998; Johnson, Booth, & O’Hearn, 2001; Meltzoff, 1995). Twelve-month-olds, however, did not show this understanding, suggesting that it emerged in the months shortly after the first birthday (Bellagamba & Tomasello, 1999).¹

But imitation is a fairly demanding response measure, and so the question arises

---

¹ Other studies of younger infants that used looking time measures (e.g., Woodward, 1998) did not test infants’ understanding of failed attempts and accidents, so, in our view, they do not provide clear evidence of understanding of others’ goals.
whether infants can demonstrate the same understanding in another task paradigm at a younger age. We thus developed a procedure to use infants’ natural behavioral responses in a social interaction to test younger infants’ understanding of others’ failed attempts and accidents. Behne, Carpenter, Call, and Tomasello (2005) engaged infants in a game in which an adult gave them toys across a table. Interspersed were trials in which the adult held up a toy but did not give it over. In some cases this was because the adult was unwilling, in various ways, and in other cases it was because she was unable in various ways, each of which involved failed attempts or accidents (e.g., she could not extract the toy from a container or she dropped it clumsily while attempting to give it to the infant). In reaction to these activities, 9-, 12-, and 18-month-olds, but not 6-month-olds, showed more signs of impatience (e.g., reaching, turning away) when the adult kept the toy for herself than when she was making a good faith effort to give it over. Infants thus were first able to infer the adult’s goal, even when it did not match what actually happened, by age 9 (but not 6) months.

A further question is when infants begin to understand not just an actor's goal, but also her intention – her plan of action for achieving that goal, including the rational basis for this choice of plan. As yet, the only studies investigating this understanding have used imitation as a response measure. In a recent study from our laboratory (Schwier, van Maanen, Carpenter, & Tomasello, in press), which was inspired by a similar study of slightly older infants by Gergely, Bekkering, and Király (2002), we found that 12-month-old infants could infer an adult demonstrator’s intention and determine why she chose the particular action she chose to accomplish her goal. Infants in both of these studies watched the adult use a particular action to achieve some end. In one condition the adult had to use that action (because an alternate, more usual action was blocked in some way), and in the other condition the adult freely chose to use that action (there was no barrier to the more usual action). Infants responded differently in the different conditions, copying the particular action demonstrated by the adult more often when the adult had freely chosen to perform that action than when she was forced by her circumstances to use it, suggesting that infants saw the adult’s action in each case as chosen for some rational reason and thus worthy (or not) of
In summary, by 9 months of age infants do not just perceive others’ surface bodily motions, but rather they go deeper and interpret others’ action as a function of their goal, seeing others as persisting past failed attempts and accidents to achieve their goals. By 12 months of age, infants in addition are beginning to understand others’ intentions, including the rudiments of the way others make rational decisions in choosing action plans for accomplishing their goals in particular reality contexts.

**Understanding goals and intentions in cultural learning**

The extent to which infants know *what* others are trying to do (their goal) and *why* they are doing it the way they are doing it (their intention) affects how deeply they can participate in cultural learning. Knowing *what* others are trying to do (their goal) is important because it allows learners to filter out goal-irrelevant aspects of a demonstration, for example accidental actions (Carpenter et al., 1998), and it focuses learners on relevant aspects of a demonstration. For instance, recently we have shown that 12-month-old infants, like older, preschool-aged children (Bekkering, Wohlschläger, & Gattis, 2000), interpret and reproduce the same adult actions differently depending on what they see as the adult’s goal (Carpenter, Call, & Tomasello, 2005). More dramatically, in some cases, knowing an adult’s goal may enable children to succeed at a task at which they could not otherwise succeed. For example, Carpenter, Call, and Tomasello (2002) demonstrated to five groups of 2-year-old children how to pull out a pin and open a box. What differed across groups was what children experienced just prior to this demonstration, with some children receiving information about the demonstrator’s goal (i.e., what she intended to do with the box as she approached it) and some not. Children who knew before the demonstration that the demonstrator intended to open the box were later significantly better at opening the box themselves than children in each of several control conditions who received no prior information or only goal-irrelevant information. Interestingly, children who did not know what the demonstrator was about to do performed just as poorly as children who received no demonstration at all.
In addition to reading others’ goals, the ability to read others’ intentions (i.e., engage in rational imitation) tells learners how the demonstrator is achieving her goal and why she is doing it in this way. This is especially important in human cultural learning, when sometimes it is necessary to do things the way others do (for example when learning the conventional use of artifacts or communicative symbols; see Gergely & Csibra, in press; Tomasello, 1999) and sometimes it is not. Once infants understand the rational dimensions of action and choice, they are better able to know when they should follow the adult and when this is not necessary, in which case they can simply pursue any means that they think is effective for themselves, as they did in the studies of Gergely et al. (2002) and Schwier et al. (in press). Being able to make this distinction is an important ability for 1-year-old infants as they begin to participate in earnest in the cultural activities around them.

**Apes’ understanding of goals and intentions**

There is now growing evidence that chimpanzees understand something about others’ goals, even at the level of failed attempts and accidents. For example, Call, Hare, Carpenter, and Tomasello (2004) tested chimpanzees in a food-giving context similar to that of the Behne et al. (2005) study with human infants described above. Similar to human 9-, 12-, and 18-month-olds, chimpanzees gestured more and left the area earlier when the human was unwilling than when he was unable – in which case they tended to wait patiently throughout his well-meaning but unsuccessful attempts. The chimpanzees apparently understood the behavior of the human in the unable conditions as persistent attempts (trying) to give them food.

In addition, three studies have tested chimpanzees’ understanding of failed attempts in a social learning context, using Meltzoff’s (1995) behavioral re-enactment procedure in which participants are shown either a successful, completed action or an unsuccessful, uncompleted attempt. Two of the studies (Myowa-Yamakoshi & Matsuzawa, 2000 and Call, Carpenter, & Tomasello, 2005) were compromised by high levels of performance in a baseline condition, but the third study (Tomasello & Carpenter, 2005) found that three young
enculturated chimpanzees, like the 18-month-old children in Meltzoff’s study, completed the action equally as often when they saw a failed attempt as when they saw the completed action.

Finally, Call and Tomasello (1998) tested apes' ability to distinguish purposeful from accidental actions in a different paradigm. They trained subjects to associate a marker situated on top of one of three opaque buckets with the location of hidden food. On test trials a human then placed the marker on one of the buckets purposefully, but either before or after this he let the marker fall accidentally onto one of the other buckets. Apes as a group chose the bucket that had been marked in a purposeful manner.

Apes thus show some understanding of others’ failed attempts and accidents – their goals. However, there is as yet no evidence that apes understand others’ intentions as rational choices of action plans. In an ongoing series of studies in our lab, (Buttelmann, Carpenter, Call, & Tomasello, in preparation), we have presented apes with many different versions of Gergely and colleagues’ (2002) study, using tasks that are more relevant to apes (involving using tools to retrieve food). The overall finding is that apes do not respond differentially depending on whether the demonstrator was forced to use the action he used or freely chose to use that action. This is the case even for a small group of chimpanzees who have been raised in a human, cultural environment (Tomasello & Carpenter, 2005).

If apes do understand something about others’ goals but not their intentions, this could help explain the general finding that in social learning situations, unlike human children, apes typically copy the end result and not the particular actions the demonstrator used (see, e.g., Call & Carpenter, 2003, for a recent review). If apes do not understand that the demonstrator chose to use those particular actions for some reason, then it is no surprise that they do not copy those actions.

Conclusions

In summary, by their first birthday, human infants understand that others have both goals and
intentions and in many cases are able to infer what those goals and intentions are. This enables infants to learn from others in special ways, filtering out goal-irrelevant aspects of demonstrations and completing failed attempts, and, of particular importance in learning cultural or conventional behaviors, knowing when it is necessary to copy the exact way a more experienced demonstrator did something. As we shall see below, understanding others’ goals, and especially their intentions, will also enable infants to participate in collaborative interactions with other members of their culture. Apes may understand others’ goals but currently there is no evidence that they understand others’ intentions. This limits the types of cultural learning they can engage in and their ability to collaborate with others (see below).

1.2 Understanding perception, attention and knowledge

Research on infants’ understanding of others has mostly focused on the understanding of goal-directed and intentional action. However, also important is infants’ understanding of what other people perceive, attend to and know. For example, when 9- to 12-month-olds begin to understand that actors strive for goals they must know also that the actor perceives and monitors her actions and pays attention to those parts of the environment which are relevant for her goal. So, what do infants understand about what others perceive, attend to and know?

Understanding perception

Infants’ understanding of perception has for the most part been investigated with the gaze following paradigm. There is fairly strong agreement that by the end of their first year of life, infants follow an adult’s gaze reliably to an outside target (e.g., Corkum & Moore, 1995, Carpenter, Nagell & Tomasello, 1998b). But more than just following an adult’s head direction, infants by around 14 months have been shown to understand some important aspects about the seeing process. This was tested with several modifications of the classic gaze following paradigm, in which the adult’s vision of the target was manipulated in various
ways. The results show that 14-month-old infants understand that i) an adult’s eyes need to be open in order for her to see an object (e.g., Brooks & Meltzoff, 2002), ii) the eyes and not just the head need to be oriented towards the object (e.g., Caron, Butler, & Brooks, 2002), and iii) the adult’s line of sight to the object needs to be unobstructed (e.g., Dunphy-Lelii & Wellman, 2004). In sum, these findings indicate that by 14 months of age, infants have an understanding of some necessary conditions for seeing. For even younger infants of 12 months of age, however, the results were ambiguous. For example, in a study by Caron, Kiel, Dayton and Butler (2002), 12-month-olds did not understand that an adult could see an object through a window in a screen, but not through an opaque barrier. Similarly, in Brooks and Meltzoff’s (2002) study, 12-month-olds did not know that an adult could not see when her eyes were covered by a blindfold. It is important to note, however, that these tasks may be particularly demanding, because infants were confronted with conflicting information, on the one hand the adult’s behaviour (turning her head in the direction of an interesting target) and on the other hand the obstruction of the adult’s view.

In order to investigate more closely what infants understand about visual perception, we thus took a different approach (Moll & Tomasello, 2004). Unlike the existing variations of the gaze following paradigm, we blocked the infant’s and not the adult’s view to a target by various kinds of barriers, to see whether infants of 12 to 18 months of age would follow an adult’s gaze to these locations. We thus created a situation in which the infant did not have immediate perceptual access to the target. Instead, the target was located outside of the infant’s immediate visual field, and she had to either crawl or walk some distance in order to be able to see what the adult saw. In each of two studies, we found that 12- and 18-month-old infants crawled or walked a short distance in order to look behind a barrier an adult was looking behind and thereby see what the adult was seeing. They did not do this in two control conditions in which a barrier was present but no one was looking behind it (the adult was looking across the room or in front of the barrier). This is strong evidence that infants in this age range understand that others see things and that they are motivated to see what others are seeing. These findings are perhaps not so surprising for 18-month-olds, as a number of
studies using a variety of different methods all converge on this conclusion (Brooks & Meltzoff, 2002; Butler et al., 2000; Corkum & Moore, 1995; Moore & Corkum, 1998). They are more surprising for 12-month-olds because previously there were few studies showing that 12-month-olds have some understanding of seeing. The results also show that Butterworth (1995) underestimated 12- to 18-month-olds’ gaze following skills. According to him, “babies’ capacity for joint attention is limited by the boundaries of the visual field […] it is as if the infant […] fails to comprehend the possibility of a space outside the range of immediate visual experience” (Butterworth & Jarrett, 1991, p. 56). The current results demonstrate that, contra Butterworth, infants’ gaze following is more than a mere geometric extrapolation of a line of sight within the immediate visual field. One-year-olds can move beyond their immediate visual field (current study) and look behind themselves (Deák, Flom, & Pick, 2000) in order to see what another person is seeing.

**Perspective taking**

However, gaze following tasks like these do not involve any perspective-taking. All the infant needs to know is *that* there is something to be seen behind the barriers. Together with a curiosity of what that something is, this suffices to be successful in these tasks, which do not require the infant to take the adult’s perspective and imagine or predict what the adult can or cannot see from her visual point of view.

Research on the development of understanding of visual perspectives has focused on toddlers and especially preschoolers of 4 to 4½ years of age (see Flavell, 1992). The tasks which have been developed for these age groups cannot be easily used with infants, because they pose relatively high linguistic and other task demands on the children. For instance, in probably the best-known task for toddlers and preschoolers, children have to respond verbally to the question of what they themselves and an adult can see (Masangkay, McCluskey, McIntyre, Sims-Knight, Vaughn, & Flavell, 1974). Other tasks involve non-verbal production measures, in which the child is asked to place an object relative to a screen such that an adult’s visual access to the object is blocked (e.g., Flavell, Shipstead, & Croft, 1978; McGuigan & Doherty, 2002). Children solve these tasks at 2½ years of age, but younger
children have simply not been tested with these measures.

We investigated younger children’s abilities to take another person’s visual perspective using a new procedure with 18- and 24-month-olds (Moll & Tomasello, in press-a). The question was whether these young children would know which of two objects was not visible from the perspective of an adult because her vision to it was blocked by an occluder. In the experimental condition, the adult ambiguously searched for a toy, whereas in the control condition, she just made a neutral request. The result was that the 24-month-olds, but not the 18-month-olds correctly handed the adult the hidden toy significantly more often in the experimental than in the control condition. Thus, by 2 years of age, children knew what the adult could and could not see and how this was related to her searching behavior. In contrast to the gaze following studies, this perspective-taking study required children to demonstrate that they know precisely what the adult could and could not see at a specific moment—without being given the opportunity to determine this on the basis of superficial cues like line of regard. The main difference between an understanding of perspectives and gaze following is thus that the child must know and be able to specify the content of what the other person sees.

It thus seems that infants can follow gaze (even to unperceived spaces) much earlier than they demonstrate an understanding of visual perspectives. One reason for this might be that following another person’s gaze direction is much simpler than determining her focus when the only information which helps disambiguate the focus is the perceptual availability of the objects to that person. But usually, adults engage with their infants in richer interactions, in which it becomes more clear what they are focused on. This is in line with the findings from some studies on understanding attention, which we will turn to now.

**Understanding attention and knowledge**

Thus far, we have looked at infants’ understanding of perception, namely, the fact that perception is object-directed (gaze-following studies) and is always bound to a specific visual point of view (visual perspective-taking). However, when we pursue a goal in acting, we do
not just perceive but also selectively attend to what we think is relevant for our action and the achievement of our goal. Attending, more than perceiving, thus involves a concentration or focus on a specific part or feature of what is perceptually present (e.g., Husserl, 2004; James, 1890). From the perspective of the outside observer, it might seem more difficult to know what a person is selectively focused on than to just determine what is in her line of regard. But there is reason to believe that infants actually do not find it harder to understand selective attention. That is, understanding attention can be seen as analogous to understanding intention: in both cases, the infant needs to know that an actor or perceiver chooses one means or aspect over another for a reason (either to enact or to attend to). And since one-year-olds already show some understanding of this when interpreting others’ intentional action (Gergely et al., 2002, Schwier et al., in press; see section 1.1), it is possible that they also understand attention to the same extent, namely, that people make decisions for reasons, in both action and perception.

A study by Tomasello and Haberl (2003) shows that indeed by one year of age, infants understand selective attention. In their study, 12- and 18-month-olds played with an adult with two toys in turn. Before a third toy was brought out by an assistant, the adult left the room. During her absence, the infant played with the third toy together with the assistant. Finally, all three toys were held in front of the infant, at which point the adult returned into the room and exclaimed excitement followed by an unspecified request for the infant to give her a toy (without indicating by gazing or pointing which specific toy she was attending to). Surprisingly, infants of both ages selected the toy the adult was attending to because it was new for her. In order to solve this task, infants had to understand 1) what the adult knew and did not know in the sense of what she had and had not become acquainted with from previous experience, and 2) the link between novelty and attention: namely, that people often attend to unknown objects.

There is, thus, converging evidence that by 12 to 14 months of age, infants have some understanding not only of perceiving, including some necessary preconditions for seeing (e.g., Caron et al., 2002), but they can also determine what others selectively attend to, even
when this requires in addition an understanding of what the person knows and does not know by acquaintance. This has now been demonstrated not just with the selection paradigm described above, but also with a different interactive response measure (Moll, Koring, Carpenter, & Tomasello, in press) and habituation measures (Onishi & Baillargeon, 2005).

**Apes’ understanding of perception and attention**

There is growing evidence that chimpanzees and other great apes understand some important aspect of others’ visual perception (see Call & Tomasello, 2005, for a review). For example, they follow the gaze direction of both conspecifics and humans to external targets (e.g., Itakura, 1996; Povinelli & Eddy, 1996; Tomasello, Call and Hare, 1998), they check back to the looker (and eventually quit looking) if nothing is there (Call, Hare & Tomasello 1998; Povinelli & Eddy 1996, Tomasello, Hare & Fogleman 2001) and they even follow the gaze direction of humans to targets behind barriers (Tomasello, Hare & Agetta, 1998). Importantly, they also show some understanding of the relation between what others can see and what they will do (Hare, Call, Agetta & Tomasello, 2000; Hare, Call & Tomasello 2001; Hare, Call and Tomasello, in press; Melis, Call & Tomasello, in press).

In contrast to the wealth of research on apes’ understanding of visual perception, very little is known about apes’ understanding of others’ attention. In the one experimental study that did address this question the chimpanzees tested (i.e., three young human-raised chimpanzees), did not demonstrate an understanding of others’ selective attention (Tomasello & Carpenter, 2005).

1.3 Conclusion

This series of studies showed that infants as young as 12 to 14 months have an understanding of some important aspects of others’ intentional action and perception. In particular, the research suggests that infants this age already understand that other act and attend to things for reason. That is, infants recognize that others act in the pursuit of goals and that they choose means to fulfill these goals given their circumstances or constraints. Similarly, they
also show some understanding that others perceive and selectively attend to certain aspects of the environment.

Importantly, recent studies have demonstrated that chimpanzees, too, understand some aspects of intentional action and perception. In particular, the new findings suggest that they read others’ behavior as goal-directed, distinguishing between intentional and unintentional actions. Furthermore, they also show an understanding of what others can and cannot see. There is no evidence, however, that non-human apes understand the choices that an actor makes in creating action plans and focusing attention. This may be one contributing factor for the differences between apes and humans regarding processes of cultural learning and creation. However, another more fundamental factor giving rise to these differences may be the motivation and ability to engage in shared activities, as will be discussed next.

2. Engaging in shared activities: from joint attention to joint pretense

Humans’ cultural skills do not only rely on understanding other persons’ actions and perception. The development of cultural practices and skills over historical and ontogenetic time also requires the ability to engage with others in joint activities with shared goals and intentions (Tomasello et al., 2005). We here examine the proposal that the ability to participate in cultural learning and creation crucially depends on skills of shared intentionality that young children already begin to develop during infancy. Thus we will first look at infants’ joint attentional skills and their role in social-cognitive development (section 2.1). Then, we will review our research on infants’ and young children’s ability to communicate with others (section 2.2), to cooperate with others (section 2.3) and to pretend with others (section 2.4). Finally, we will focus briefly on children with autism to examine whether the relative strengths and impairments of children with autism may be related to specific deficits with respect to their motivation and skill to engage in shared activities (section 2.5).
2.1 Joint attention and engagement

Infants in their first year of life interact with both their social and their physical environment. That is, a 6-month-old infant may interact dyadically with objects, grasping and manipulating them, or she may also interact dyadically with other people, expressing emotions back-and-forth in a turn-taking sequence. But at around 9 to 12 months of age a qualitatively new set of behaviours begins to emerge that are triadic in the sense that they involve a referential triangle of child, adult and object/event to which they share attention. Infants now jointly engage and share attention with others to third entities or events, they follow others’ focus of attention and they also actively direct others’ attention using pointing gestures (see Carpenter et al. 1998b, for details).

Importantly, these joint-attentional skills open up the possibility to participate in processes of cultural learning and creation. For example, a number of studies have demonstrated the role of joint attentional engagement for toddlers’ language acquisition (see Tomasello, 1999, for a review). The Vygotskian proposal we wish to examine here is that joint engagement with others also scaffolds young children’s social cognitive development. Specifically, we address the hypothesis that joint engagement is a powerful way for very young infants to develop an understanding of other people, in particular to come to know what others have and have not experienced.

The role of joint engagement in understanding attention and knowledge

In some situations infants as young as 12 months are already able to determine what others are selectively attending to, based on the other person’s previous experience (Tomasello & Haberl, 2003, see above). That is, when an adult showed excitement and then made an unspecified request for one of three toys, infants gave the adult the toy that was new to her, even though for the infants themselves all three toys were equally familiar. How do infants at this young age do this? In order to investigate how infants come to know what others attend to based on what they know, we conducted a series of studies using Tomasello and Haberl’s
basic procedure (2003). Specifically, we explored the conditions under which infants know which of several objects is unknown for an adult and thus catches her attention.

In the first study (Moll & Tomasello, in press-b), infants observed an adult experience each of two objects (known objects) and then leave the room while infants played with a third object (unknown object, target) along with an assistant. The adult experienced the first two objects in one of three different ways. The first was a situation of joint engagement in which infant and adult played with the objects, looking at and manipulating them together (Joint Engagement condition). The second was a situation of individual engagement in which the infant observed the adult actively manipulating, inspecting, and reacting to the objects by herself (Individual Engagement condition). The third was a situation of onlooking in which the infant observed the adult simply watching as the infant played with the objects (Onlooking condition). We tested 14- and 18-month-olds in these three conditions, using as a dependent measure the object infants selected when the adult returned to the room and exclaimed excitement. The results showed that infants of both age groups knew which objects the adult knew in the Joint Engagement condition, only the 18-month-olds knew this in the Individual Engagement condition, and infants at neither age knew this in the Onlooking condition. These results suggest that infants are first able to determine what adults know on the basis of their direct, triadic engagements with them. Only a few months later, by 18 months, infants are less dependent on these joint engagement contexts, as is evidenced also by word learning studies (Baldwin, 1993; Floor & Akhtar, 2006).

To zero in on the joint engagement effect found for the younger infants, we conducted two follow-up studies with 14-month-olds (Moll, Carpenter, & Tomasello, under revision) and looked more closely at i) whether joint engagement also helps infants understand what others know from visual experience alone, which generally seems hard for them (as we know from the Onlooking condition, see above), and ii) whether it is sufficient for these young infants to witness the adult interact with a third person around the objects in order to register her as being acquainted with them. The results showed that, as long as the adult reacted in joint engagement as she watched the infants playing with the objects (without manipulating
them herself), infants recognized which objects were known for her. However, infants did not
distinguish between the known and the unknown objects when they witnessed, from a third
person perspective, the adult jointly engage with another person around the known objects, in
a situation similar to ‘overhearing’ word-learning studies (see Akhtar, 2005).

**Conclusion**

By 12- to 14 months of age, infants can determine what others selectively attend to, even
when this requires in addition an understanding of what another person knows and does not
know by acquaintance. The current studies show that this kind of social-cognitive ability
develops inside joint engagement episodes, in which the focus of the partner somehow
becomes especially transparent and is perhaps made mutually manifest (Eilan, 2005; Heal,
2005). This might be because in these contexts, infants form with their partner a shared focus
or even a shared goal (Tomasello et al., 2005). In any case, just like joint engagement seems
to help infants determine the referents for words (e.g., Dunham, Dunham, & Curwin, 1993;
Tomasello & Farrar, 1986), and engage in more mature forms of play (Bigelow, MacLean, &
Proctor, 2004), it also helps them very generally to monitor and understand what others
attend to and know.

**2.2 Communicating with others**

*Understanding communicative intentions*

A hallmark of human cultural learning and creation is the ability to communicate using
linguistic symbols. As socio-pragmatic accounts of language acquisition have emphasized,
acquiring the conventional use of linguistic symbols presupposes the ability to recognise the
communicative intentions expressed by others, and this relies on infants’ understanding of
intentional action and their ability to participate in joint attentional scenes (Bruner, 1983;
Tomasello, 1999). Importantly, however, understanding communicative acts presents a
special case of intentional and attentional understanding. One aspect that distinguishes
communicative intentions from intentions simpliciter is that to understand your communicative intention, I must recognise your intention towards my attentional state (e.g., Tomasello, 1999). Many analysts have pointed out that people read others’ communicative intentions based on the assumption that the communicator directs the addressee’s attention in ways that are relevant to their joint interaction or background (e.g., Sperber & Wilson, 1986). The question, here, is whether infants already understand others’ communicative acts in this way, too.

One-year-old infants reliably follow others’ gaze direction and pointing gestures. However, this does not necessarily mean that they recognise others’ communicative intentions. A better situation for assessing infant understanding of communicative intent is one in which the infant follows an adult gesture to an otherwise uninteresting target and, in addition, needs to recognise why the adult took the trouble to direct her attention to that target. Following Sperber and Wilson (1986), we must look for situations in which the infant asks herself: Why did the adult do this for me? Why is this object to which he is gesturing relevant to our interaction?

One possible task with this structure is the so-called object choice task that has been employed with both human children and non-human primates (see, e.g., Povinelli, Reaux, Bierschwale, Allain, & Simon, 1997; Tomasello, Call & Gluckman, 1997). In this task, an adult hides a reward in one of several opaque containers and then indicates the reward’s location by giving a communicative cue, for example pointing to the baited container. This research showed that children aged 2.5 and 3.0 years not only followed the adult's indication to one of the containers, but they also inferred that the hidden reward could be found there (as evidenced by their search behaviour). The children treated each communicative attempt as an expression of the adult’s intention to direct their attention in ways relevant to the current interaction/game. In contrast, in the same situation great apes did not infer the location of the hidden food. This was not because they cannot follow the direction of pointing or gazing (they can; see Call & Tomasello, 2005, for a review), but because they did not tune in to the adult's communicative intention and infer why he was directing their attention to this
They did not understand that the gesture was made for their benefit, and so they did not seek or find the relevance of this act in this context.

Whereas the ability of human children aged 2.5 and 3 years has been demonstrated in a number of investigations, it was unclear whether younger children – at the age when language first begins to emerge – were also able to understand others’ communicative intentions in this way. Thus, to address this issue, we engaged one-year-olds in hiding-finding games, which were based on the object-choice task. In particular, we were interested in whether children this age can recognize the communicative intent behind non-linguistic gestures (a) when infants were addressed directly and (b) when they ‘overheard’ communicative gestures directed at a third person.

Firstly, we examined infants’ understanding of communicative gestures when they were addressed directly. Thus, after an introduction to the hiding game, an adult hid a toy in one of two identical opaque containers and then, addressing the infant, indicated the toy’s location by giving a communicative cue – either pointing or ostensive gazing toward the correct location. Children aged 14, 18 and 24 months participated in this game. At all three ages they reliably searched in the correct container, indicating that they were using the adult’s communicative cues to find the hidden toy. A control study demonstrated that infants’ successful search performance was not simply based on low-level gaze following mechanisms. That is, when the adult produced similar surface behaviour as before, but without expressing communicative intent (e.g., looking absent-mindedly at the baited container), infants’ search performance was at chance level (Behne, Carpenter & Tomasello, 2005). Taken together these findings suggest that infants as young as 14 months are able to recognise the communicative relevance of ostensive behaviour directed at them.

Secondly, we examined whether infants can also understand the communicative intent of ostensive behaviour that is not directed at them but instead at a third person (i.e., when infants are “overhearing” others’ communicative interactions). Therefore, the same hiding-finding game was set up as described above, but throughout the game the adult directed her communicative cues (pointing and ostensive gazing) at another adult, without addressing the
When searching for the hidden toy, 18-month-old infants (and to a certain extent 14-month-olds, too) reliably chose the correct container, indicating that they recognised the communicative relevance of the gestures that they were “overhearing” (Gräfenhain, Behne, Carpenter & Tomasello, submitted).

**Conclusion**

In contrast to the poor performance generally shown by non-human primates in this type of task (see Call, Agnetta, & Tomasello, 2000 for a review), human infants used others’ communicative gestures to guide their search, both when being addressed directly and when “overhearing” gestures directed at others. Importantly, this required more than low-level gaze or point following abilities. For successful search performance infants needed to attend to the communicative cue, recognise its relevance, and identify its referent. Our findings suggest that infants as young as 14 months are able to recognise other persons’ communicative intent within the frame of an ongoing activity. As discussed above, this ability is essential for the acquisition of language, as well as for other forms of cultural learning and shared co-operative activities.

**Infant pointing**

One-year-olds not only respond to attempts by others to direct their attention, they also actively direct others’ attention themselves using pointing gestures. Infant communicative pointing has been proposed to express two distinct performatives (e.g., Bates, Camaioni & Volterra, 1975; see also Camaioni, 1993): proto-imperatives are requests for objects and proto-declaratives are attempts to share attention and interest to objects or events. These gesture appear to be truly communicative acts – as evidenced by the fact that while gesturing infants often alternate gaze between the adult’s face and the object, suggesting that they check whether the adult is following and responding to their gesture (Bates, et al. 1975; see Franco & Butterworth, 1996, for an elaboration).
However, some researchers have expressed skepticism that 12-month-olds’ produce declarative gestures to share attention and interest (Moore & Corkum, 1994) or that their declarative points are even communicative at all (Desrochers et al. 1995). For example, Moore and colleagues argued that early points typically identified as protodeclaratives were in reality aimed at gaining adult attention to the self, rather than sharing attention and interest with the other person (Moore and Corkum, 1994; Moore & D’Entremont 2001). To address this controversy, we conducted a series of detailed studies on infant declarative pointing. Specifically, we investigated the following three questions: Firstly, do infants intend to communicate when they point declaratively? Secondly, do infants want to direct others’ attention specifically to what they point at? Thirdly, what are the underlying motives for infant pointing?

**Declarative pointing**

To investigate infant declarative pointing we have used a paradigm in which pointing is elicited on a number of trials by interesting events, like puppets appearing or lights flashing from behind a large screen at a distance. On each trial, if the infant points, an experimenter (E) responds consistently in one of several ways in different experimental groups. We then measure the infants’ reactions to the experimenter’s responses to determine which response might satisfy infants’ intent when pointing.

In a first study we tested four hypotheses about what 12-month-olds want when they point (Liszkowski, Carpenter, Henning, Striano & Tomasello, 2004). To test Desrochers et al.’s (1995) hypothesis that infants point non-communicatively for themselves, E neither attended to the infant nor to the indicated event in response to infants’ points (Ignore condition). To test Moore and D’Entremont’s (2001) hypothesis that infants do not want to direct or share attention and just want to obtain attention to themselves, E never looked at the event and instead attended to the infant’s face and emoted positively to the infant (Face condition). To test the hypothesis that infants only want to direct attention and nothing else, E only attended to the event (Event condition). And to test our hypothesis that infants want
to share attention and interest, E responded to the infant’s point by alternating gaze between the event and the infant, emoting positively to the infant about the event (Joint Attention condition).

The main findings were that infants pointed communicatively to share attention and interest with E about the event. Infants preferred the Joint Attention condition and pointed on significantly more trials in that condition compared to each of the other three conditions. When E attended only to the infant but never to the event and emoted positively (Face condition), infants repeated their pointing within a trial to that event significantly more often than in the Joint Attention condition, in an apparent attempt to direct E’s attention to it. Thus, contrary to the account of Moore and D’Entremont (2001), infants wanted to direct E’s attention to an event. However, when E only attended to the event and did not look and comment back to the infant (Event condition), infants also repeated their pointing within a trial and, in addition, looked significantly more often to E than in any other condition. Just directing attention without sharing was thus not satisfactory either. Instead, infants expected some sort of a comment from E, indicative of sharing interest in the event that both attended to.

In a second study (Liszkowski, Carpenter & Tomasello, submitted) we followed up on these results and investigated to which extent infant pointing already involves two main components of linguistic speech acts: reference and attitude (see Searle, 1969). Using the same general paradigm as in the previous study, E either understood infants’ referent and attended to what they pointed at, or he misunderstood infants’ referent and instead attended to a barrier which obstructed his line of sight to infants’ referent. In addition, we manipulated whether E emoted positively, showing interest, or neutrally, showing disinterest, controlling the attitude expression about a referent. The main findings were that 12-month-olds preferred the condition in which both referent and attitude were shared (Joint Attention), pointing on significantly more trials in that condition than in each of the other three. When E emoted positively but misunderstood infants’ referent (Misunderstanding condition), infants were not satisfied and repeated pointing within a trial significantly more often than in the Joint
Attention condition, apparently to re-direct the adult’s attention to the correct referent. When E attended to the referent but was not interested in it (Uninterested condition), infants pointed overall on significantly less trials, but did not repeat their pointing within a trial – not simply requesting positive emotions imperatively.

These two studies demonstrate that 12-month-olds point with the intent to communicate and want to direct others’ attention specifically to what they point at. Pointing in the current situation is motivated by sharing an attitude about the referent with a communicative partner. In sum, findings show that infant pointing is a joint communicative act, to comment on and point out something for the other person and to share attitudes about it with the other person. Infants’ declarative pointing is thus not motivated egocentrically to obtain something for self. Instead, it is motivated by sharing and aligning self and other in some way.

*Informative pointing*

Interestingly, adults also point for other reasons than sharing interest. For example, adults also point to the lost keys that someone is looking for, helping her find them by providing necessary information. The motive of such informative pointing is different from that of declarative pointing. In informative pointing the providing of information is not to engage mutually about the referent but instead mainly to benefit the other person. The pointing is more about the recipient’s relation to the referent than the sender’s, to help her find an object. Helping is an important feature of human shared cooperative activities (see Bratman, 1992) but has previously not been investigated in infants this young.

In a third study (Liszkowski, Carpenter, Striano, & Tomasello, 2006) we therefore investigated whether 12-month-olds also point like adults to help others find what they are looking for by informing them of something they do not know. We designed a search paradigm, in which E first repeatedly demonstrated an action with one of two objects. Both objects then disappeared in various ways (e.g., they fell down or somebody put them away) and E then began searching ambiguously. Findings were that infants pointed at the object E
was looking for, significantly more than at the other simultaneously misplaced object, and without requestive accompaniments like whining, reaching or repeated pointing, apparently to inform E of its location. This study thus showed that infants point to provide information for others which requires the cognitive ability to detect what information is relevant for an adult. In addition, it revealed for the first time in infants this young the presence of a prosocial motive of helping others freely and without direct benefit for self.

**Conclusion**

It has been proposed that declarative pointing at 12 months is not communicative (Desrochers et al., 1995) or that it only serves to obtain attention to the self egocentrically, without any social-cognitive understanding of others’ attentional states (Moore & D’Entremont, 2001). Our studies on infant pointing do not support such lean views. Instead, we have presented evidence that 12-month-olds point communicatively with the goal of directing others’ attentional states, and with cooperative, prosocial motives such as sharing interest in things and helping by providing information for others.

Further studies in our laboratory are investigating when infants point at things others are already attending to, when infants point out what is new to others, how infants refer to absent referents, whether infants know that the addressee has to be able to see their pointing, and what role caregivers’ pointing and infants’ cognitive development play in the emergence of pointing. All findings converge on a new view of infant pointing as a prelinguistic fully communicative act which involves both reference and attitude, pro-social motives like sharing and helping, and a social-cognitive understanding of others’ psychological perspectives and experiences (Liszkowski, 2005; 2006, for an overview).

Interestingly, there is no evidence of any of this in the gestures of non-human primates. That is, there is no evidence that non-human apes produce declarative gestures or that they freely provide information for others that is not related to their own needs (Tomasello, in press, Tomasello & Carpenter, 2005, Haimerl et al. in preparation). Thus, infants’ declarative and informative pointing seems to be a uniquely human form of
communication, reflective of their motivation and social-cognitive skills to engage in shared activities.

2.3 Cooperating with others

Social behaviors such as helping and cooperation are interesting both cognitively and motivationally: In order to successfully help another person with a problem, the helper must understand the other’s unachieved goal and possess the altruistic motivation to act on behalf of the other. Such altruistic motivations are rare evolutionarily: In fact, several researchers have claimed that only humans act altruistically towards non-kin - in contrast to other primates who strive only after their own best benefit (Alexander, 1987; Fehr & Fischbacher, 2003), regardless of the benefits or costs for others (Jensen, Hare, Call & Tomasello, 2006; Silk et al., 2005).

Whereas for helping, understanding the other’s individual goal might be sufficient, cooperative activities are based upon the formation of a shared goal, including the motivation to mutually support each other’s actions to reach that goal (Bratman 1992, Tomasello et al., 2005). Cooperative activities are thus another test case for the proposal that humans – and maybe only humans – engage in social interactions which are characterized by shared intentionality.

To investigate the phylogenetic and ontogenetic roots of these behaviors, we conducted a series of comparative studies. Here, we present two studies in which human children and chimpanzees were tested on a similar set of helping and cooperation tasks. Such comparisons are intriguing because they may enable us to distinguish behaviors which were present already in our common phylogenetic ancestor from aspects which are unique to the human lineage.

Helping in young children and chimpanzees

A number of studies have demonstrated that young children show concern for others in distress – as a kind of emotional helping (see Eisenberg & Fabes, 1998, for an overview).
Also, as Liszkowski et al. (2006) have shown (see above), one-year-old infants help others by informing them about the location of an object they are looking for. However, no experimental studies have systematically assessed young children’s instrumental helping – providing help to people who are faced with a problem and are unable to solve it on their own.

We thus designed a study aimed at investigating such instances of instrumental helping in young children at 18 months of age (Warneken & Tomasello, 2006). To do so, we developed ten situations in which an experimenter was performing a specific action, but suddenly encountered a problem and needed assistance from the child to achieve his goal. It turned out that children performed spontaneous, unrewarded helping behaviors in diverse situations: For example, they helped the adult retrieve an out-of-reach object like a pen he had accidentally dropped on the floor and was unsuccessfully reaching for; they completed his stacking of books after his failed attempt to do this, or they opened the door of a cabinet when his hands were full. Importantly, the children did not perform these actions in control conditions in which no help was necessary (e.g. when he had thrown the pen on the floor on purpose), showing that their behavior depended upon the other’s goal.

Interestingly, when we gave the same ten (slightly varied tasks) to three human-raised chimpanzees (aged 36 to 54 months), they also helped in one type of situation: They helped a human caregiver by handing her objects she was unsuccessfully reaching for, but they did not help in the other kinds of situations (e.g. completing an action, opening a door for the other). The findings indicated that the chimpanzees were able to understand goal-directed action, at least when the goal was easy to discern like in situations where a person is reaching for an object. Moreover, they demonstrated the altruistic motivation to instrumentally help the other, questioning the assumption that altruism is unique to humans. Ongoing research in our laboratory is attempting to clarify (1) whether this is restricted to the close relationship between the nursery-raised chimpanzees and their caregiver and (2) whether it extends beyond this context, which may resemble an imperative request for handing over objects.

Taken together, these results show that young children seem to have a natural tendency to help others in a variety of ways – even when the other person is a stranger and
children receive no benefit for helping. Furthermore, it appears that the common ancestor of chimpanzees and humans already possessed rudimentary forms of skills and motivations to help others.

Cooperative activities in young children and chimpanzees

We then may ask whether the same is true for activities that are based upon a shared goal. Engaging in cooperative activities with a shared goal is a crucial achievement in human ontogeny as it enables partners to reach goals which lie beyond the means of an individual, forming the basis for cultural creation. At what age do young children begin to engage in such activities by successfully coordinating their own actions with those of a partner to work towards a shared goal? Are chimpanzees motivated and able to engage in such activities as well?

To investigate these questions, Warneken, Chen, and Tomasello (in press) tested children at 18 and 24 months of age on four different cooperation tasks and compared their performance with that of the same three human-raised chimpanzees mentioned above. The rationale of all these tasks was that they could not be performed successfully by one person alone but required the joint activity of two people. Two of these tasks were cooperative games (such as a “trampoline” in which two people had to hold a large piece of cloth and make a toy bounce on it) and two others were cooperative problem solving tasks (e.g. two people had to perform complementary roles like one person holding a container open so that the other person could retrieve an object from inside). Children at both ages were able to successfully cooperate with an adult partner in most of both types of tasks (games and problem solving). In addition, we found that children at 24 months were coordinating their actions with the partner more skillfully than children at 18 months. For example, the older children would position themselves in the correct location more quickly and adjust their actions to their partner temporally by holding the container open until the other had completed his action. The most interesting findings were obtained when the adult partner interrupted his participation at a predetermined moment (see Ross & Lollis, 1983 who first
devised this method): Children of both age groups frequently communicated to the partner in an attempt to request his cooperation. All children produced at least one such communicative attempt: They frequently used gestural communication like pointing to the apparatus or placing the apparatus in front of the partner, and at 24 months they often accompanied this with verbalizations. This at least shows that the children understood their own and their partner’s action as interconnected parts of a joint activity. Moreover, this can also be taken as evidence that the children were trying to redirect their partner towards a shared goal, perhaps insisting on the commitment to support each other’s actions in a cooperative activity.

When the three human-raised chimpanzees (aged 33 to 51 months) interacted with a human caregiver in the same tasks, they were able to solve a problem-solving task such as lifting a door so that the partner could retrieve a piece of food from inside (see also Melis, Hare & Tomasello, 2006). In contrast, they showed no interest in co-operating in the social games, such as the trampoline game which had no other goal beyond doing something together. Most importantly, they never once attempted to reengage their partner when she refrained from cooperating – even in the problem-solving task in which they were able to successfully coordinate their action. The chimpanzees instead tried to solve the task alone, or else disengaged from the task completely, suggesting that they had not formed a shared goal with the other.

Human beings appear to be especially adapted for cooperative activities. Our chimpanzees – even though being raised in a human environment – did not develop similar skills, at least with regard to forming shared goals with others. These findings therefore provide evidence for a uniquely human form of cooperative activity involving shared intentionality, which emerges in the second year of life.

2.4 Pretending with Others

Young children from the second half of their second year not only engage with others in shared cooperative activities of instrumental and playful kinds – they also enter into the collective space of fictional worlds in pretend play. This type of play is a uniquely human
phenomenon (see Gomez & Martin-Andrade, 2005). Part of the reason this is so, we think, is not so much that other species lack imagination, but that pretending is a form of social cooperative activity that is acquired by means of cultural learning, and is founded on an ability to participate in collective intentionality, both of which are absent in other species.

Cultural learning of pretense

The debate about the origins of pretending in ontogeny has been shaped by two contrasting approaches, Piagetian Individualism and Vygotskian culturalism. According to Piaget (1962), the child first creates pretend play autonomously, through individual rather than social processes and through interactions with the environment rather than with people. In contrast, the Vygotskian Soviet school has considered play as essentially situated in specific social and cultural contexts (see El’Konin, 1966). In other words, whereas Piaget thinks children invent their toys individually first, the Vygotskian tradition claims toys become toys initially in the same way that tools become tools – by cultural learning. In line with the Vygotskian tradition (see also Tomasello 1999; Tomasello et al., 1993), we hypothesized that in fact pretend play is acquired as are other action forms, namely by cultural imitative learning. We pursued this claim in a recent set of studies on children’s acquisition of pretense (Rakoczy, Tomasello & Striano, 2005). The design was intended to simulate what might be called the cultural ontogeny of artefacts: young children’s first encounters with hitherto unknown and (for them) functionless objects and their subsequent learning of how to use those objects as either tools or toys. Young children (18 and 24 months old) were first shown a series of novel objects. An adult then demonstrated instrumental actions on some and pretense actions on others. In a second phase the infants were then given the objects and could act with them several times themselves. The results were as follows: (1) Children imitated both kinds of actions in similar ways, with the same object as the demonstrator (though imitation rates were lower in absolute terms for pretense acts, and 18-month-olds were almost at floor in imitating pretense), (2) Children produced hardly any creative pretense acts (but many creative instrumental acts), and (3) During pretense acts children showed significantly more frequent
and stronger forms of social behaviour, namely gazing (and in one study smiling) to the adult. These results thus suggest that children in their second year begin to imitate pretense actions with objects in similar ways to their imitation of simpler kinds of actions. They also suggest that early pretense is only minimally creative and is an essentially social activity, in which creativity and solitary forms develop later. Tools become tools for children in similar ways as toys become toys—through picking up the intentional and cultural affordances and functions of objects by observing adults’ actions with these objects.

However, pretend play, we think, is not acquired on the basis of blind mimicry, but instead is understood by young children as a meaningful, non-serious activity. This understanding of pretending as a just another, albeit special kind of intentional activity enables children to imitatively acquire pretend play action forms, and to enter into shared cooperative pretense. Shared pretense, in fact, can be considered one of the earliest areas in which children enter into collective intentionality: Not only does the child observe others perform individual pretense acts of the form “I pretend X” and then performs such an individual “I pretend X” act as well; also, the child joins into shared cooperative “We pretend X” acts together with others. More specifically, this involves a mutual commitment to act together within the game as indicated by the production of actions that respect the implications of each other’s pretend stipulations. For instance, if you pretend to ‘pour’ water, I can pretend to ‘drink’ it and must be careful not to ‘spill’ it (see Rakoczy, in press, for details of this idea).

This view of pretend play, however, is a rich one and stands in contrast to a prominent view of early pretense understanding in the recent theory of mind literature. This is the so-called ‘behaving-as-if’ construal (e.g. Lillard, 1993, 1998; Nichols & Stich, 2000; Perner, Baker & Hutton, 1994). The basic claim of the ‘behaving-as-if’ theory is that young children up to the age of four to five years do not yet understand the intentional structure of pretending, but merely have a very superficial understanding of pretending as a somehow deviant type of behaviour (‘behaving-as-if’), without grasping that pretending is intentionally and non-seriously acting-as-if.
Understanding pretense intentions

In another set of studies, we set out to pit our claim and the ‘behaving-as-if’ theory against each other (Rakoczy, Tomasello, & Striano, 2004; Rakoczy & Tomasello, in press). The logic of the studies was straightforward: children’s imitative and inferential responses to two kinds of as-if-behaviours—pretending and trying—were compared. Two- and three-year-old children were shown pairs of superficially analogous incomplete as-if-behaviours with objects: pretending to do an action and unsuccessfully trying to do the same action, for example to pour from a (closed) container into a cup. In both cases the actor made pouring movements with a novel container over a cup, but without actually pouring. In one case, he marked the activity with signs of playfulness and sound effects as pretending to pour, and in the other case he marked it with signs of surprise and frustration as trying to really pour. Importantly, the container really did contain water and thus could really be used to pour. In the first study the situation was set up as an imitation game. After the actor’s demonstration children were given the object and could perform actions with it themselves. Three-year-olds (and to a weaker degree 2-year-olds) very clearly showed that they understood pretending and trying as such: after trying models, they performed the complete action themselves (e.g. actually poured the water) or tried to really perform it, often commenting on their failure (e.g. “I cannot do it either”). But after pretense models, they instead pretended themselves without appearing to care about the real effects of their acts (e.g., whether there was water coming out of the container).

In another study, children were presented with some of the same demonstration pairs, but not in a strict imitation game. Rather, the pragmatics of the situation was set up to encourage more creative inferential responses. In this study we found that when the 2- and 3-year-olds saw an actor try to pour, they themselves then really did the action or tried to, but did so using different means. For example, they used a tool to open the container first. When the actor had pretended to pour, in contrast, children themselves pretended to drink and to give a Teddy bear a drink. Thus, children showed a rich understanding of the intentional
structures of pretending and trying as different forms of behaving-as-if: They grasped that in trying to pour, the actor wanted to perform the action properly and intended to make the proposition “there is water coming out of this container” true by bringing it about. They also understood that, in pretending to pour, the actor was acting intentionally but non-seriously as if pouring and as if the proposition “there is water coming out of this container” was true. Accordingly, these two kinds of behaviours license very different inferences that children grasped: in the trying case, other means should be used when one wants to perform the same action, but now successfully. In the pretense case, perceiving the other person’s pretending as such (e.g., pretending to pour into a cup) licenses joining into a shared pretense scenario (“We pretend that there is water in the cup now”).

In summary, these results suggest that young children by two years of age imitatively acquire pretend play from others, understand the basic intentional structure of pretending and, based on this understanding, join into inferentially structured joint pretense founded upon a collective intention of the form ‘We pretend…’.

**Talking about pretending**

Note that in the studies reported so far, children’s understanding of others’ pretending and their ability to join into pretense were investigated as indicated in their systematic and inferential actions. In another study (Rakoczy, Tomasello, & Striano, in press), we directly compared this understanding in action to children’s explicit understanding in words – the first study in the area of pretense comprehension to look at different levels of understanding with one and the same task. The same action demonstrations as in the “pretending-trying” studies were presented to children at 3, 4 and 6 years of age, and children were asked whether the character pretended to perform an action (e.g., drink) or tried to perform this action. This explicit verbal task proved rather difficult for children up to the age of six. That is, there was a huge décalage between understanding implicit in children’s systematic, appropriate and inferential responses at 2 years, and their explicit understanding as expressed in words at 6 years.
How is such a huge décalage to be explained? Again, in line with the Vygotskian tradition and recent cultural learning approach (Tomasello 1999), the idea is that of a dialectical development: Children’s social understanding enables them to enter into shared activities, to learn culturally (i.e., imitatively and collaboratively), and to acquire social practices, particularly language and discourse. Participation in specific explicit forms of discourse in turn enables new levels of reflective understanding, social and otherwise (Tomasello, 1999). In the specific case of pretense understanding, specific pretense discourse of the form “We pretend to X” and “We pretend that p” etc. should thus be crucial in developing a later explicit understanding of pretending (as measured in our verbal “pretending-trying” tasks and in many other verbal pretense understanding tasks). Support for this central role of explicit pretense discourse comes from a recent training study we conducted (Rakoczy et al., in press). Two groups of 3.5-year-old children received intensive experience with diverse pretense activities (and were compared with a control group that received functional play experience). For the Explicit Group, the pretense experience was accompanied by explicit discourse making use of “pretend that” (e.g. “I’m pretending that this stone is an apple, but really it is a stone”) and “pretend to” (e.g. “She is pretending to give him an apple”) constructions. In the Implicit Group, in contrast, the pretense scenarios were talked about implicitly, making use of specific implicit pretense discourse markers (e.g., “This is my apple” in a funny voice). In the post-test, only the Explicit Group showed improvement on pretense-reality distinction tests, in which children had to verbally state what an actor had pretended of an object and contrast this with what it really was (Flavell et al., 1987), and tests in which children had to tell whether someone had pretended or tried to do an action verbally. Explicit discourse about pretending and the pretend identities of objects thus turned out to be crucial in the development of reflective understanding of pretense.

**Conclusion**

Pretend play is a form of collective intentionality, involving the joint creation and respect of status functions. These are functions which only exist because they are collectively assigned
to objects and which are expressed in the formula “X counts as Y in context C”: “This piece of paper counts as money in our currency area” or “This piece of wood counts as the queen in chess” are the standard examples (Searle, 1995). The collective non-literal treatment of objects in pretend play, we argue, can be seen as the ontogenetic cradle of such status function creation which is at the heart of uniquely human institutional reality: “This (wooden block) counts as an apple in our pretense game” (Walton, 1990).

Since, according to our studies, young children from the age of 2 years participate in pretend play as a prominent form of collective intentionality, the question remains of exactly how much they understand about status functions. Two very basic characteristics of status function assignment are the following (Searle, 1995). First, status functions are context-relative: Only in context C does X count as a Y, but not in other contexts. For example, a piece of wood counts as a queen in chess but not in other board games. Second, status functions carry normative force: “X counts as Y in context C” implies that X ought to be treated as a Y in context C and ought not to be treated in other ways. For example, a dollar bill ought to be treated as money and not only as a piece of paper in the relevant currency area, and a queen in chess should be treated as having certain powers and not only as a piece of wood.

Such qualities as normativity and context-relativity are what Vygotsky observed in pretend play when he proposed that “Whenever there is an imaginary situation in play, there are rules” (1978, p. 95). In our current research we are investigating children’s understanding of these aspects of collective status function creation. In a recent study, for example, we found evidence that young 3-year-olds have a nascent understanding that one object can simultaneously count as different objects in different pretense contexts (Wyman, Rakoczy & Tomasello, in preparation). Regarding the normative dimension of status function creation, in a pilot study we found that 3-year-olds actively protested against pretense acts by a play partner that were normatively inappropriate given the joint status function assignment (i.e., not treating the X as a Y in the relevant context) (Rakoczy, unpublished data).
In sum, children from 2 years of age participate in joint pretending, a form of collective intentionality which involves the joint assignment of status functions. As status functions are the conceptual building blocks of institutional reality, children’s joint games of make believe can be considered an ontogenetic cradle for entering into institutional reality.

2.5 Children with autism

Autism is a neurodevelopmental disorder that is characterized by qualitative impairments in social interaction and in communication. Several theoretical proposals have been put forward, attempting to explain the pattern of abilities and difficulties shown by children with autism, among them the proposal that the core feature of autism is a problem in “theory of mind” understanding (see e.g., Baron-Cohen, 2000 and Frith 1997 for reviews). Importantly, recent research suggests that children with autism do show some understanding of other people, in particular with respect to some aspects of intentional action (Aldridge, Stone, Sweeney, & Bower, 2000; Carpenter, Pennington & Rogers, 2001) and perception (Leekam, Baron-Cohen, Perrett, Milders, & Brown, 1997). Thus, the question arises whether the relative strengths and impairments of children with autism may be accounted for by specific deficits related to their motivation and ability to engage in joint activities with shared goals (see also Hobson, 2002). In order to assess this proposal we briefly review the ability of children with autism with respect to engagement in the shared activities discussed above, that is, joint engagement, communication, pretend play and co-operation.

Firstly, deficits with regard to joint attention are so pervasive in children with autism that they actually represent diagnostic criteria (DSM-IV, American Psychiatric Association 1994). Children with autism show very little coordinated joint engagement and they rarely initiate joint attention with others by declaratively pointing to or showing objects (e.g., see Baron-Cohen 1989, Charman et al. 1997). They also rarely respond to others’ bids of joint attention (see Leekam et al. 1997). Secondly, linguistic communication and the use of symbols is another problem area for children with autism, and their impaired ability to signal non-comprehension and make appropriate repairs to their own linguistic messages to help
others are well documented – suggesting that their communication is not fully collaborative (Loveland et al. 1990, Sabbagh 1999). Thirdly, with regard to pretend play, deficits in this domain are so characteristic for young children with autism that this, too, is considered a diagnostic criterion.

Less is known, however, about the co-operative abilities of children with autism. It has been reported that children with autism engage in relatively little cooperative play with peers (Lord 1984), but interacting with peers may be more demanding than engaging in co-operative activities with an adult, who may initially scaffold the interaction. Therefore, we explored the performance of children with autism in interactions with an adult, looking at both their helping behaviour and their engagement in co-operative activities.

To study helping, Liebal, Colombi, Rogers, Warneken, and Tomasello (submitted) presented young children with autism and matched children with other developmental delays four tasks involving an adult unsuccessfully reaching for an object, similar to the tasks Warneken and Tomasello (2006) used to elicit helping behaviors. We found that children with autism tended to help less frequently than children with other developmental delays when the adult unsuccessfully reached for an object.

To investigate cooperation, Liebal et al. (submitted) adapted the study by Warneken et al. (in press) and presented the same children who participated in the helping study with tasks that encouraged them to cooperate with an adult partner in either problem-solving tasks or social games. We found that children with autism were less likely to cooperate with an adult partner and less coordinated in their actions. This was especially the case for the social games in which the goal was not to retrieve an object (as in the problem-solving tasks) but “simply” to play the game with another person, like in the trampoline task where the only goal was to jointly bounce the block for the fun of it. During interruption periods in which the adult stopped interacting, children with autism showed fewer partner-oriented behaviors like waiting or redirecting the partner to the task, and they made fewer communicative attempts that involved eye contact than did children with other developmental delays. In sum, these findings support the proposal that children with autism show characteristic impairments with
respect to their motivation and ability to engage in social interactions that involve a form of shared intentionality.

3. Shared intentionality

Together these empirical findings provide broad support for the theoretical proposals of Tomasello et al. (2005). They suggested that we should think of the development of human social-cognition as comprising two ontogenetic strands: (1) the understanding of intentional action and perception, and (2) the ability to understand and participate in social interactions involving one or another form of shared intentionality.

The first strand of human social-cognitive ontogeny is the general primate (or perhaps only ape) understanding of intentional action. Based on the data reviewed here, as well as other data on primate cognition, we may propose that all apes, including humans, understand individual intentional action in terms of the pursuit goals, as well as the basics of visual perception. There have been a number of proposals to the effect that this skill is a hard-wired and modular part of the human perceptual system. Just as humans automatically see certain perceptual sequences as causal (Michotte 1963; Leslie 1984), they automatically see certain actions performed by animate agents as goal-directed. Gergely and Csibra (2003), for example, have proposed that human infants possess an action interpretation system that perceives human-like action as teleologically directed to a goal from the second half of the first year of life; independently developing is a reference interpretation system concerned with following gaze and the like (Csibra, 2003). Baron-Cohen (1995) proposes something similar, with two early developing innate modules involving the perceiving of goals and eye gaze direction. Soon after the first birthday an independent "shared attention mechanism", emerges, taking outputs from the two earlier modules as inputs.

Although our view shares some features with these views, there are two important differences. First, we do not see infants' understanding of goals/intentions and perception/attention as blocked off from one another in a modular fashion. Indeed, much recent evidence on infant social-cognitive development suggests that in attempting to
understand what others are doing and why they are doing it, infants comprehend intentional action and perception as an integrated system (i.e., as a kind of control system). They display such an integrated understanding from 9 months of age when they know that an actor pursues goals persistently (until he perceives that the world matches his goal) and also engage with other persons triadically around external objects - where they must infer people's perceptions from their goals and their goals from their perceptions. In general, we do not see how an observer can understand goal-directed action (much less rational action) without understanding a perceiving organism who monitors the world for signs of success, failure, obstacles, and so forth.

Second, we believe that to understand the origins of a human cognitive skill we must go beyond simply labeling it as "innate". Indeed, although we concur that understanding actions as goal-directed is a biological adaptation, this says nothing about the ontogenetic process. It is very unlikely, in our view, that a human or ape kept in social isolation for the first year of life would suddenly understand others as goal-directed or intentional agents on its initial encounter with them; presumably the developmental pathway for understanding intentional action depends on species-typical social interactions early in ontogeny. This does not necessarily mean, however, any specific experiences. Thus, Kaye (1982) proposes that to understand intentions infants must themselves be treated by adults as intentional, in the sense that adults interpret their actions in adult-like terms and provide various types of feedback to this effect. The problem with this more specific hypothesis is that there seems to be fairly wide cultural variation in how infants are treated by adults - with adults in some cultures not really treating infants as fully intentional - and, by all accounts, all children in all cultures develop an understanding of others as intentional agents.

The second strand of human social-cognitive development is the sharing strand. Theorists such as Trevarthen (1979), Stern (1985), Braten (2000), and especially Hobson (2002), have elaborated the interpersonal and emotional dimensions of early human ontogeny in great detail. We mostly agree with their accounts, but we find that they do not give sufficient attention to the other, intention-reading, line of social-cognitive development. Our
proposal is that the uniquely human aspects of social cognition emerge only as uniquely human social motivations interact with an emerging, primate-general understanding of animate and goal-directed action - which then transforms the general ape line of understanding intentional action into the modern human line of shared intentionality.

Although the precise nature of this interaction is not entirely clear, our general view is that infants begin to understand particular kinds of intentional and mental states in others only after they have experienced them first in their own activity and then used their own experience to simulate that of others (Tomasello 1999; see Sommerville & Woodward, 2005, for experimental evidence supporting this view). However, contrary to our previous view, we do not think that simple "identification with others" is a sufficient basis for the simulation process - certainly not if we mean bodily identification, as there is now evidence that neonatal chimpanzees engage in the same kind of facial mimicking as human infants (Myowa 1996; Myowa, Tomonaga, & Matsusawa, 2004), and even some species of birds are good at copying actions (e.g., Zentall 1996). And so we would speculate at this point that more deeply psychological levels of identification with others - of a kind sufficient to enable individuals to simulate the intentional and mental states of others on analogy with their own - depends crucially on the skills and motivations for interpersonal and emotional dyadic sharing characteristic of human infants and their caregivers (Hobson, 2002).

Again one can imagine that a species-typical social environment, involving human-typical social interactions with other persons, is required for the emergence of the sharing motivation and its related skills of social engagement. But again some theorists have proposed that some kinds of specific experiences are necessary. For instance, Stern (1985) proposes that parents must "mirror" back to infants their own emotions or behaviors and Gergely (2003) posits an especially important role for certain kinds of social contingencies in terms of timing. But again it is not clear that children in all cultures receive such experiences, or that children who are deprived of them end up unable to share psychological states with others. And so the ontogenetic process for sharing emotions and intentions with others may be fairly robust in the face of different particular human social environments.
Our proposal for the early developmental pathway characteristic of human social cognition is thus that it is the synergistic product of the general ape line of understanding intentional action, unfolding from 9 to 14 months (based on earlier recognition of object-directed actions, see Woodward, 1998), and the modern human motivation to share psychological states with others, present from very early in human ontogeny. There is almost no research establishing a solid relationship between any kind of particular social experience infants might have and individual differences in the unfolding of this developmental pathway. In the absence of such studies, we might tentatively conclude that this is a very robust, heavily canalized ontogenetic pathway in humans that (in the absence of neurobiological disorders) emerges in all 'normal' human environments.

What results from this developmental process, early in the second year of life, is the ability to engage in a number of different kinds of collaborative interactions involving shared intentionality. As a part of this, there also emerges a new form of cognitive representation, what Tomasello et al. (2005) called dialogic cognitive representations in which each participant conceives the collaboration holistically, with both roles in a single representational format (see also Fernyhough, 1996). Although we have no concrete evidence for where these novel forms of cognitive representation come from, our supposition is that they are Vygotskian internalizations of these special collaborative activities. These representations then enable children's participation in cultural (mediated) practices such as linguistic communication and other forms of symbolic interaction. Dialogic cognitive representations thus include and go beyond theoretical constructs such as "identification with others" (Hobson 1993; Tomasello 1999b), the "like me" stance (Meltzoff & Gopnik 1993), and "self-other equivalence" (Barresi & Moore 1996) - which may be ontogenetic forerunners. That is to say, they capture the fact that the child both knows that she is in some sense equivalent to others - actors can substitute for one another in acts of imitation and role reversal - but at the same time she is different from others. Dialogic cognitive representations thus have built into them the functional equivalence (though not identity) of different participants in activities, one of whom may be the self, but they have additional aspects (e.g.,
intentions about the other's intentions) deriving from the motivation to share psychological states with others.

In any case, our research efforts will continue to be aimed at answering questions focused on the process by which human beings have evolved the skills and motivations to create and participate in the cultural activities and practices that constitute human societies.

References


Gräfenhain, M., Behne, T., Carpenter, M., & Tomasello, M. (submitted). One-year-olds’ understanding of nonverbally expressed communicative intentions directed to a third person.


Moll, H., Carpenter, M., & Tomasello, M. (under revision). Fourteen-month-old infants know what others know only from joint engagement.


