

# Chapter 16

## The Gap is Social: Human Shared Intentionality and Culture

Michael Tomasello and Henrike Moll

**Abstract** Human beings share many cognitive skills with their nearest primate relatives, especially those for dealing with the physical world of objects (and categories and quantities of objects) in space and their causal interrelations. But humans are, in addition, biologically adapted for cultural life in ways that other primates are not. Specifically, humans have evolved unique motivations and cognitive skills for understanding other persons as cooperative agents with whom one can share emotions, experience, and collaborative actions (shared intentionality). These motivations and skills first emerge in human ontogeny at around one year of age, as infants begin to participate with other persons in various kinds of collaborative and joint attentional activities. Participation in such activities leads humans to construct during ontogeny, perspectival and dialogical cognitive representations.

### 16.1 Introduction

The gap we are trying to explain is obvious: we humans live in complex societies dependent on complex technologies, symbol systems, and social institutions – whereas other primate species do not. We, humans, are scientifically investigating and writing about them, not they about us.

Following Vygotsky (1978) and Tomasello (1999), the general proposal here is that the human gap is best explained in terms of, ultimately, social (or cultural) factors. That is, human beings are especially sophisticated cognitively not because of their greater individual brainpower, but rather because of their unique ability to put their individual brainpowers together to create cultural practices, artifacts, and institutions – underlain by skills and motivations for shared intentionality – which

---

M. Tomasello (✉) and H. Moll  
Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany  
e-mail: tomasello@eva.mpg.de, moll@eva.mpg.de

are then passed along to youngsters as a second line of inheritance in the species, resulting in a ratcheting up of cultural and cognitive complexity over historical time. A child raised alone on a desert island, or even by chimpanzees, would cognitively not be very different from the apes, as its unique adaptation for absorbing culture would be intact but there would be nothing there to absorb.

But it turns out that identifying the cognitive and social-cognitive factors that enable developing human beings to take advantage of the cultural practices, artifacts, and institutions around them is not so easy. The approach in our research group for some years has been to focus on situations and phenomena that are mostly simpler than those in modern adult life, but still have all of the key characteristics for exploring the gap between human and ape cognition. Toward this end, we have focused on young children as representative of the human species, with the idea that they help us to abstract away from the complexities of adult human life and get down to the essentials. For comparison, we have focused on great apes, especially chimpanzees, as humans' closest primate relatives.

What we will do here is three things. First, we will present evidence that the gap is indeed social, drawn especially from a recent large-scale study of the full range of cognitive skills in great apes and human children. Second, we will review some recent studies specifying in more detail the nature of the difference between apes and human children in situations involving shared intentionality in (1) collaboration, (2) communication, and (3) social learning. Finally, we will offer some speculations as to how these small-scale cooperative abilities of shared intentionality scale up into uniquely human cognition and culture.

## 16.2 Human and Great Ape Cognitive Skills Compared

An obvious hypothesis about the human difference is that human beings simply have bigger brains and so more “general intelligence” than other animals: more memory capacity, greater inferential skills, faster learning, further foresight and planning, finer skills of perceptual discrimination – more and better of everything. And this quantitative difference somehow translates into a qualitative difference in cognitive abilities.

In a recent study, we tested this hypothesis – in comparison with what we called the cultural intelligence hypothesis – by giving a very large test battery (a kind of nonverbal IQ test) to two of humans' closest primate relatives, chimpanzees and orangutans, along with 2-year-old human children (Herrmann et al. 2007). If humans simply have more cerebral computing power and general intelligence, then the children should have differed from the apes uniformly across the different types of tasks. But that was not the result. The result was that the three species were very similar when it came to cognitive skills for dealing with the physical world – problems having to do with space, quantities, and causality – but the human children were much better than the two ape species when it came to cognitive skills for dealing with the social world – problems of imitative learning, gestural

communication, and reading intentions.<sup>1</sup> When correlational analyses were done, no general intelligence, or *g*-factor, was found (Herrmann et al. in press). These findings are not compatible with the hypothesis that the cognitive difference between humans and apes is a simple function of more cerebral horsepower.

There have also been a few hypotheses about more specific computational capacities that might make human cognition unique – though they cannot explain these cross-species test battery results either. In one hypothesis, for example, humans are better able to perform multiple tasks simultaneously. In another, humans are better with relational categories and making analogies across different materials. These hypotheses may have some validity, but it is difficult to see how multitasking alone or analogy-making could account for the results of our cross-species study of intelligence, much less such things as linguistic symbols, social institutions, and cultural norms. These are all collective cultural products that are not easily accounted for by simply adding up the computational power of individuals either generally or in specific skills.

Uniquely, human cognitive skills are not simply the result of greater computational power overall, or of some increase in a specialized cognitive ability. Rather, they result from an ability enabling humans to put their heads together, so to speak, in cooperating and communicating with one another in ways that led to the creation of complex cultural products, including both material and symbolic artifacts, such as linguistic symbols. Human children grow up in the midst of these material and symbolic artifacts, and by learning to use them in interaction with others (as well as internalizing these interactions cognitively), they actually create, during ontogeny, evolutionarily new ways of thinking. This hypothesis sets the problem that must be addressed if we want to provide a plausible evolutionary account of how human cognition became, in effect, a collective enterprise.

### 16.3 Cultural Activities in Humans and Great Apes

Humans' adaptation for living and exchanging information in cultural groups manifests itself in many ways. What we would like to focus on here is the three main classes: (1) small-scale collaborative activities (translating ultimately into large-scale social and cultural institutions); (2) cooperative communication (transforming ultimately into language); and (3) cultural learning (resulting ultimately in cumulative cultural evolution).

---

<sup>1</sup>One might object that the human subjects had an advantage over the non-human ones in the social tasks, because they interacted with a conspecific vs. a member of a different species. However: (1) many of the tasks were chosen based on previous research showing no difference in performance as a function of conspecific or human interactants (e.g., ape vs. human demonstrators in social learning); and (2) the comfort level of all subjects in the testing situation was assessed and, if anything, the children were most shy (and this measure did not correlate with overall cognitive performance in the main tasks).

### 16.3.1 Collaboration

Individuals of virtually all primate species engage in group activities on a daily basis. These activities may be considered cooperative in the sense that they coordinate their behavior temporally and spatially with the other animals in the group. However, as in previous theoretical work (Tomasello et al. 2005), here, we want to single out for special attention “shared cooperative activities” – a subtype in which humans routinely engage. In our modified version of Bratman’s (1992) characterization, joint or shared cooperative activities are mainly characterized by the three features: (1) the participants in the cooperative activity share a joint goal, to which they are jointly committed; (2) the participants take reciprocal or complementary roles in order to achieve this joint goal; and (3) the participants are generally motivated and willing to help one another to accomplish their role, if needed.

*Joint commitment to a goal.* One group activity that has been posited as being especially complex is chimpanzee group hunting. Boesch and colleagues (Boesch and Boesch 1989; Boesch and Boesch-Achermann 2000; Boesch 2005) have observed chimpanzees in the Tai forest hunting in groups for arboreal prey, mainly monkeys. In the account of these researchers, the animals take complementary roles in their hunting. One individual, called the driver, chases the prey in a certain direction, while others, the so-called blockers, climb the trees and prevent the prey from changing directions. An ambusher then silently moves in front of the prey, making an escape impossible. Of course, when the hunting event is described with this vocabulary of complementary roles, then it appears to be a joint cooperative activity: complementary roles already imply that there is a joint goal, shared by the role-takers. But the question really is whether this vocabulary is appropriate at all. A more plausible characterization of the hunting event, from our perspective, is as follows: each animal fills whatever spatial position is still available at any given time so that the encircling is accomplished in a stepwise fashion, without any kind of prior plan or agreement to a shared goal or assignment of roles. Then, without pursuing a joint goal or accomplishing a certain role within a higher-order framework, each individual chases the prey from its own position (see also Moll and Tomasello 2007a). This event clearly is a group activity or group action, because the chimpanzees are “mutually responsive” as they coordinate their behaviors with that of the others in space and time (see also Melis et al. 2006). But what seems to be missing is the “togetherness” or “jointness” that distinguishes shared cooperative activities from other sorts of group actions.

This interpretation is strongly supported by studies that have investigated chimpanzees’ abilities to cooperate in experimental settings. In one study, Warneken et al. (2006) tested three juvenile human-raised chimpanzees with a set of four different cooperation tasks. In two of these tasks, a human tried to engage the chimpanzee to cooperate in order to solve a problem (e.g., extracting a piece of food from an apparatus). In the other two tasks, the human tried to engage the ape to play a social game. The authors looked at two things: the chimpanzees’ level of behavioral coordination and the chimpanzees’ behaviors in the so-called

interruption periods, in which the human suddenly stopped participating in the activity. The results were very consistent: in the problem-solving tasks, chimpanzees coordinated their behaviors quite well with that of the human, as shown by the fact that they were mostly successful in bringing about the desired result, as, for instance, extracting the piece of food from the apparatus. However, they showed no interest in the social games, and so the level of coordination in these tasks was low or absent. Most important was what happened when the human suddenly interrupted the activity. In none of the tasks did a chimpanzee ever make a communicative attempt to reengage the partner. Such attempts were missing even in cases in which they should have been highly motivated to obtain the desired result, as in the problem-solving task involving food. The absence of any efforts by the chimpanzees to reengage their human partner is crucial: it shows that the chimpanzees did not cooperate in the true sense, since they had not formed a joint goal with the human. If they had been committed to a joint goal, then we would expect them, at least in some instances, to persist in trying to bring it about and in trying to keep the collaboration going.

For humans, the situation is different from very early on in ontogeny. Warneken et al. (2006) conducted an analogous study with 18- and 24-month-old human children. Unlike the chimpanzees, children cooperated quite successfully and enthusiastically not only in the problem-solving tasks, but also in the social games. For example, these infants enjoyed playing a “trampoline” game together, in which both partners had to simultaneously lift up their sides of a small trampoline with their hands, such that a ball could bounce on it without falling off. Most importantly, when the adult stopped participating at a certain point during the activity, every child at least once produced a communicative attempt in order to reengage him. In some cases, the children grabbed the adult by his arm and drew him to the apparatus. The older children of 24 months of age also often made linguistic attempts to tell the recalcitrant partner to continue. Unlike the chimpanzees, we thus find in human infants the ability to cooperate with joint commitment to a shared goal: the children “reminded” the recalcitrant partner of their shared goal and expected him to continue in order to achieve it. There was even some evidence that the children already understood the normativity behind the social games and the way they “ought to be played.” For example, in one of the games, they always used a can in order to catch a toy when it came falling out of one end of a tube after their partner had thrown it in from the other end. They could have also caught it with their hands, but they preferred to do it the way it had been demonstrated to them. This implies that they perceived the can as a constitutive element of the game, and they wanted to play the game the way it “ought” to be played. The chimpanzees, on the other hand, never used the can in order to catch the toy – if they engaged in the game at all, they simply used their hands. It, thus, seems that human infants by the age of 18 months, in contrast to apes, are able to jointly commit to a shared goal.

*Role reversal.* The second criterion for cooperation, as we define it, is role-taking. True cooperation should involve that the partners perform reciprocal roles and also understand them, in the sense that they coordinate their actions and

intentions with the possibility of reversing roles. This form of role-taking would suggest that each partner represents the entire collaboration, its shared goal and reciprocal roles, holistically from a “bird’s eye view” instead of just from within whatever role they happen to be taking at the moment. One study purporting to show role reversal in chimpanzees is that of Povinelli et al. (1992). In that study, chimpanzees were trained in one of two roles of a cooperative hiding game with a human. Some chimpanzees were trained in the role of a communicator, who indicated to the human where a piece of food was located. The other chimpanzees were trained in the complementary role of the “operator,” who extracted the food from the location indicated by the human. When the chimpanzees had learned their initial role to criterion, a role switch was initiated and the question was whether the chimpanzees would spontaneously reverse the roles. One of the chimpanzees, whose initial role was that of the communicator, was immediately successful as operator after the switch. But the problem is that this individual most likely comprehended human indicating gestures before the study – as this animal had extensive interactions with humans. The two individuals that switched to be a communicator also seemed to reverse the roles effectively, as they were reported to provide the human with cues about the location of the food fairly quickly. However, the problem in this case is that it is not clear that the chimpanzees actually produced any communicative signals at all, but instead the humans simply interpreted their natural bodily orientation to the food.

A better controlled investigation of role-reversal skills in chimpanzees was done by Tomasello and Carpenter (2005) with the same three young human-raised chimpanzees which participated in Warneken et al.’s (2006) study. In this study, a human demonstrated to the chimpanzee various actions with each of four pairs of objects. For each pair of objects, one functioned as a “base” and the other as an “actor.” The human then demonstrated to the chimpanzee how the two, the actor and the base, are put together. For instance, she put a “Tigger” figure on a plate and “Winnie the Pooh” figure in a little toy car. Then E gave the actor (e.g., Tigger) to the chimpanzee and held out the base (the plate) towards the chimpanzee, thus offering that the chimpanzee put the actor on the base to complete the act. If chimpanzees did not perform the role of putting the actor on the base spontaneously, E encouraged them to do so by vocalizing and, and if they still did not respond, by helping them put the actor on the base. To test for role reversal, E then handed the chimpanzee the base and held out the actor to see whether she would spontaneously offer the base. Two of the three chimpanzees held out the base object at some point. But, crucially, none of these responses occurred spontaneously, and more importantly, in none of these responses was the holding out of the base accompanied by a look to E’s face. A look to the partner’s face while holding out the object is a key criterion of “offering” used in all studies with human infants (Bates 1979; Camaioni 1993). Thus, in Tomasello and Carpenter’s (2005) study, there was no indication that the chimpanzees offered the base to the human, and so there were no acts of role reversal.

An analogous study with human infants of 12 and 18 months of age was conducted by Carpenter et al. (2005). As in the study with the chimpanzees,

situations were set up in which an adult did things like hold out a basket in which the infant was asked to place a toy. After the infant complied, in the test for role reversal, the adult placed the basket within the infant's reach and held up the toy herself. Impressively, even some of the 12-month olds spontaneously held out the basket for the adult while at the same time looking to her face, presumably in an anticipation of her placing the toy inside. Thus, the infant's handing behaviors, in contrast to those of the chimpanzees, were clearly the acts of offering learned through role reversal. It, thus looks as though chimpanzees, in contrast to young human children, do not fulfill either of the first two criteria of cooperation: sharing a joint goal and understanding the roles of a joint activity in some general way.

*Mutual support.* The third criterion is that, if needed, the partners of a joint cooperative activity help one another do their part successfully. This criterion is inherently linked to the other two: the commitment to the joint (and not just an individual) goal implies a responsibility not just for the successful completion of one's own role, but also to some degree for that of the other participants, and so helping them fulfill their goal (or, in some instances even replacing them) is an integral part of true collaboration. In two recent studies, chimpanzees did not take an opportunity to "help" another individual to obtain food (Silk et al. 2005; Jensen et al. 2006). But food is a resource over which apes are used to compete, and so helping might be better investigated in situations that do not revolve around food. Given our interest in helping as a constituent of collaboration, the most important form of helping is "instrumental helping," in which one individual helps another instrumentally to achieve a behavioral goal. We know of only one study investigating instrumental helping in nonhuman primates. Warneken and Tomasello (2006) had three human-raised juvenile chimpanzees watch a human attempt, but fail to achieve, different kinds of individual goals. Reasons for her failure were that her desired objects were out of reach, that she ran into physical obstacles, clumsily produced wrong results, or used ineffective means. The chimpanzees helped the human in some cases. However, the range of situations in which they helped was very limited: only when the adult effortfully reached but failed to grasp objects did the chimpanzees help by fetching them for her.

An analogous study was conducted with 18-month-old human infants, who also saw an adult fail to reach her goals for the same reasons (Warneken and Tomasello 2006). In this study, infants as young as 18 months of age helped the adult in various scenarios: for instance, they spontaneously removed physical obstacles that hindered the adult (e.g., they opened a cabinet so that the adult could place books inside) and showed him means that they knew were effective to bring about the intended result. It thus seems that, even though some helping behavior can be found in nonhuman primates, only human infants display helping actions in a variety of situations, providing whatever help is needed in the given situation.

What we conclude from these experimental studies is that, despite their group hunting in the wild, chimpanzees do not have "we-intentionality" (see Bratman 1992; Searle 1995; Tuomela 2002). They do not form a joint commitment to a shared goal and they do not perform reciprocal roles in the true sense as they do not generally understand roles from a bird's eye view, in the same representational

format. Finally, they seem to be limited in their abilities to help another individual – which is a necessary prerequisite to engage in cooperative activities narrowly defined. Human infants and young children, in contrast, have this we-intentionality and act cooperatively from at least 14 to 18 months of age. They “remind” their partner of the joint commitment to a shared goal, as they reengage her when she suddenly interrupts the activity (Warneken et al. 2006; Warneken and Tomasello 2007); they begin to reverse and understand their roles as early as 12 months of age (Carpenter et al. 2005); and they help others in the fulfillment of their individual roles in various ways by at least 14–18 months (Warneken and Tomasello 2006, 2007).

### ***16.3.2 Cooperative Communication***

A related domain, which also requires some form of cooperation is communication. As noted above, chimpanzees usually perform poorly in experiments that require some understanding of cooperative communication. Here, we address this issue in more detail by first looking at nonhuman primates’ own production of communicative gestures, and then at their comprehension of such gestures produced by others.

Chimpanzees gesture to one another in different contexts. Some of these gestures are clearly intentional, in the sense that they are not just triggered by certain environmental conditions, but used flexibly to do such things as elicit play in the other (by an “arm-raise”) or to request nursing (by a “touch-side”). That these gestures are indeed used flexibly is illustrated by a number of phenomena, for instance, the fact that visual gestures are only used in instances in which the recipient is visually oriented towards the sender (e.g., Tomasello et al. 1997a; Kaminski et al. 2004). One might think that if chimpanzees can gesture flexibly and understand some things about visual perception (see Call and Tomasello 2008), they should also use gestures to direct another chimpanzee’s attention to a certain event or object by pointing. There are certainly occasions in which it would be very helpful if one ape pointed for another ape to indicate the locus of some relevant event. It must, therefore, seem somewhat surprising that, in fact, there has not been a single reliable documentation of any scientist in any part of the world of one ape pointing for another. But captive apes which have had regular interactions with humans point for their human caretakers in some situations. Leavens and Hopkins (1998, 2005) conducted a study with chimpanzees in which a human experimenter placed a piece of food outside of the ape’s reach and then left. When another human came in, the chimpanzees pointed to the food so that the human would get it for him (pointing was usually done with the whole hand, but some points were produced with just the index finger; see also Leavens et al. 2004). Human-raised chimpanzees have also been found to point to humans in order to obtain access to locations where there is food (Savage-Rumbaugh 1990), and some orangutans point for humans to the location where they can find a hidden tool, which they will then hopefully use to obtain food for the orangutans (Call and Tomasello 1994).



We thus find that apes do sometimes point for humans – given that they have had some contact with humans in the past. Importantly though, they use this manual gesture imperatively only. That is, they point for humans either in order to obtain a desirable object from them directly, as in the studies by Leavens and Hopkins (1998, 2005), or indirectly by requesting from the human to provide the necessary conditions for them to get the object themselves, as in Savage-Rumbaugh's (1990) study. It, thus seems that what the apes have learned from their experience with humans is that the human will help them, and that they can use the pointing gesture instrumentally in order to make him help them. They, thus, "use" the human as a "social tool" in order to get things they otherwise could not get, and they have learned that pointing gets this tool to work (the term "social tool" was first used by Bates et al. 1975). However, no ape has ever been observed to point for another ape or for a human declaratively – that is, just for the sake of sharing attention to some entity or event, or to inform others cooperatively, as humans often do.

Liszkowski et al. (2004, 2006) have shown in a series of experiments that even when they first begin to point at around 1 year of age, human infants do this with a full range of different motives – including the motive to share attention and interest. In one study (Liszkowski et al. 2004), an adult reacted differently towards infants' points, and the infants' responses to the adult reaction were investigated. The main finding was that if the adult did not jointly attend to the event with the infant (by alternating gaze between infant and event and commenting on it) – but instead either (1) just "registered" the event without sharing it with the infant or (2) only looked and emoted positively to the infant while ignoring the event – the infants were dissatisfied and tried to correct the situation. In contrast, in the joint attention condition, infants appeared satisfied with the response. Using the same basic methodology, Liszkowski et al. (2006) found that beyond the classic distinction of imperative and declarative pointing, 12-month-olds point for others also to inform them about things that are relevant for them. In that study, they directed an adult's attention to the location of an object for which that person was searching. What this suggests is that in human ontogeny, pointing is used from the very beginning not just in order to obtain certain objects via helpful adults as social tools, but with the motivation to help/inform others or to just jointly attend to things in the world with them.

The question is thus why apes do not point to share interest and inform others as human infants do from very early in development (see also Tomasello 2006). They clearly have the necessary motor abilities to do so. And again, it would surely be useful if they spatially indicated important events for one another. So why do they not do it? To answer this question, one needs to look at apes' understanding of pointing. One of the main paradigms that has been used to assess chimpanzees' comprehension of pointing is the Object Choice task. In the task designed by Tomasello et al. (1997b), one human, the hider, hides a piece of food for the ape in one of several containers. Then another human, the helper, shows the ape where it is by tilting the container so that she can look inside and see the food. After this "warm-up," the hider again places a piece of food in one of the containers, but now

the helper indicates the location of the food for the ape by pointing at the baited container with his index finger (or by gazing at it). Variations of this method involve other kinds of communicative cues (Call and Tomasello 2005) and a trained chimpanzee instead of a human as the provider of the cue (Itakura et al. 1999). The results were the same in all these studies: the apes performed poorly, that is, they chose the correct container at chance level. They often followed the human's point (or gaze cue) to the container with their eyes, but they did not make any inferences from there about the location of food. That is, they cannot use or exploit the information that is conveyed to them via the pointing gesture – they do not know what it means. When following the human's point with their gaze, all they perceive is a useless bucket. To understand that the point is not directed at the bucket as such, but at the bucket qua location or qua container of a desired object, the apes would need to understand something about cooperation or communication. They would need to understand that the other is trying to communicate to them something that might be relevant for the achievement of their goal. In other words, an understanding of the meaning of the pointing gesture presupposes a more general understanding that others might want to help or inform us about the things which they assume are relevant for our purposes. And this understanding obviously goes beyond the apes' social-cognitive skills.

The view that the challenge of the Object Choice task does, indeed, lie in its cooperative structure is supported by recent studies using a competitive version of the task. In one version, Hare and Tomasello (2004), instead of pointing to the baited container, reached unsuccessfully for it. Superficially, this reaching behavior is very similar to the pointing gesture: the human's hand is oriented towards the container in which the food is hidden (the difference being that when pointing, only the index finger is stretched out, whereas in the case of reaching, all fingers point at the container). However, the chimpanzees' response in the reaching version was very different, as they successfully retrieved the food from the correct container. The reason for this must be that, even though the two tasks are superficially highly similar, their underlying structure is very different. Our interpretation is that in the case of reaching, the chimpanzees just need to perceive the goal-directedness of the human's reaching action and 'see' that there must be something desirable in the container. This task can, thus be solved with some understanding of the individual intentionality of the reaching action. In contrast, to understand pointing, the subject needs to understand more than the individual goal-directed behavior. She needs to understand that by pointing towards a location, the other individual attempts to communicate to her where a desired object is located; that the other tries to inform her about something that is relevant for her. So the ape would need to understand something about this directedness towards itself 'this is for me!' and about the communicative intention behind the gesture in order to profit from it. Apparently, apes do not understand that the cue is "for them" – used by the other in a helpful, informative and communicative way. Even though they are quite skillful in understanding intentional behavior that is directed at objects in the world (see Tomasello et al. 2005, for a review), they do not understand communicative intentions, which are intentions that are not directed at things or behaviors but at another individual's

intentional states (with the embedded structure: “I intend for you to know that I intend for you x”).

In order to explain why the apes fail to understand communicative intentions, one needs to broaden the perspective and focus on what we call the “joint attentional frame.” The joint attentional frame or common ground (Clark and Brennan 1991) is what gives a pointing gesture its meaning – it is what “grounds” the communication in the shared space of meaning. To illustrate the point, imagine you are walking down the aisle of a hardware store and all of a sudden a stranger looks at you and points to a bucket standing in one of the shelves. You see the bucket, but, with a quizzical look on your face, look back at the stranger, because you do not know what is going on. The reason why you do not know what is going on is that you lack a joint attentional frame with the stranger, which would give the point its meaning. The pointing as such, in this frameless scenario, does not mean anything. But if, instead, you are walking down the same aisle with a friend because you are looking for a bucket to use for cleaning purposes, and your friend points out the bucket to you, you would know immediately what he means: “Here is one!” The presence of the joint attentional frame, which could be described by something like “we are searching for a bucket,” grounds the point in the ongoing activity and gives it its meaning. Another possible scenario could be that you and your friend are looking for anything that is made of a certain kind of plastic because you like it so much. In this case, your friend’s point would have a different meaning, namely something like: “Here is an item which is made of that plastic that you like so much!” The referent of the pointing gesture thus varies as a function of the joint attentional frame in which the pointing is anchored. One can imagine an endless number of joint attentional frames for the same basic scenario, with the referents of the pointing gesture being, for instance, “item with texture of kind x,” “item which is similar to that other item we just saw,” and so forth. The pointing gesture does not just indicate some spatial location, but instead it already contains a certain perspective from which the indicated object or location is to be viewed. And the perspective is carried by the joint attentional frame.

Humans can read pointing gestures based on joint attentional frames from as early as 14 months of age. Behne et al. (2005) found that 14-month-olds choose the correct container in the Object Choice task significantly above chance, thus demonstrating that they understand the pointing gesture cooperatively. Infants also know that the validity of a joint attentional frame is limited to those people who share it. Liebal et al. (2009) had 18-month-old infants clean up with an adult by picking up toys and putting them in a basket. At one point, the adult stopped and pointed to a ring toy, which infants then picked up and placed in the basket, presumably to help clean up. However, when the adult pointed to this same toy in this same way but in a different context, infants did not pick up the ring toy and put it in the basket; specifically, when the infant and adult were engaged in stacking ring toys on a post, children ignored the basket and brought the ring toy back to stack it on the post. The crucial point is that in both conditions the adult pointed to the same toy in the same way (and everything else in the room was the same), but the infant extracted a different meaning in the two cases – based on the two different joint attentional

frames involved, and the jointness is, indeed, crucial here. Thus, in a control condition, the infant and adult cleaned up exactly as in the shared clean-up condition, but then a second adult who had not shared this context entered the room and pointed towards the ring toy in exactly the same way as the first adult in the other two conditions. In this case, infants did not put the toy away into the basket, presumably because the second adult had not shared the cleaning context with them. Rather, because they had no shared frame with this adult, they seemed most often to interpret the new adult's point as a simple invitation to note and share attention to the toy.

We, thus find that apes “communicate” individualistically, to get others to do things, and without joint attentional frames to ground the communicative intentions in a preexisting space of shared meaning. Human infants from as early as 14 months of age, on the other hand, communicate cooperatively – to share interest in things and inform others of things – and they construct and participate in joint attentional frames, which give cooperative gestures their meaning. Without a foundation in cooperative communication of this type, human language is not even thinkable (Tomasello 2008).

### ***16.3.3 Cultural Learning***

Human behavioral traditions have a cumulative history, with some of them showing a kind of “ratchet effect” of accumulating complexity over time (Tomasello et al. 1993). There is no convincing demonstration of the ratchet effect or any other form of cumulative cultural evolution for chimpanzees or any other nonhuman animals.

The explanation for this difference involved four components. First, although chimpanzees learn much socially (see for example the recent work of Whiten and colleagues as summarized in Whiten *in press*, this volume), humans seem to be more focused on actions than are chimpanzees, who are mainly focused on outcomes and goals. Humans are better and more accurate social learners: they are cultural learners. This special focus on actions enables them to socially learn activities from others in a much more accurate fashion, which not only contributes to the ratchet effect over time, but also enables the acquisition of cultural conventions that are only arbitrarily related to any causal relations in the world (such as linguistic symbols), since in this case faithful copying of actions is required.

Second, humans rely on teaching as a complement to their natural skills of social and cultural learning. Gergely and Csibra (2006) have recently elaborated an account explaining why the existence of relatively “opaque” cultural conventions (there is no causal structure or else it is difficult to see this structure) requires that human adults be specifically adapted for pedagogy toward children and human children be specifically adapted for recognizing when adults are being pedagogical (what Tomasello et al. 1993, called “instructed learning”). Engaging with others in this way is a kind of shared intentionality relying on cooperative and communication, in which the learner trusts the information given by the teacher. There has been

no systematic study of chimpanzees engaged in anything resembling teaching since the observations of Boesch (1991), which have multiple interpretations.

Third, humans imitate one another not simply when they are aimed at acquiring more effective behavioral strategies in instrumental situations, but they also imitate for purely social reasons – to be like others. The tendency of human beings to follow fads and fashions and to conform are well known and well documented, and the proposal here, following Carpenter (2006), is that this represents a different and important motivation for social learning that may produce qualitatively different behaviors. For example, human infants have a greater tendency than do chimpanzees for copying the unnecessary “style” of an instrumental action (Carpenter and Tomasello, unpubl. data), and in acquiring linguistic conventions, children are not just driven by communicative efficacy but also by the desire to do it the way the others do it (Tomasello 2003). This analysis would also explain why children in the studies cited above sometimes imitated poor demonstrators when it would have been to their advantage to ignore them, and, in general, why children copy the actual actions of others more readily than do other apes. This so-called “social function” of imitation (Uzgiris 1981), the urge to be like others, is clearly an important part of human culture and cultural transmission.

Finally, human culture persists and has the character it has, not just because human children do what others do, but also because adults expect and even demand that they behave in certain ways: children understand that this is not just the way that something *is* done, but rather the way it *should* be done. This normative judgment is another aspect of shared intentionality, as it is essentially a judgment based on the perspective of the group – how “we” do things. In a recent study, Rakoczy et al. (2008) found that 3-year-old human children not only copied the way that others did things, but when they observed a third party doing them in some other way they objected and told them they were doing it “wrong” – that is not how “one” does it. Kelemen (1999) has also shown that young children learn very quickly that a particular artifact is for a particular function, and other uses of it may be considered “wrong” – this is not how “we” use this artifact. This normative dimension to human cultural traditions serves to guarantee their faithful transmission across generations in a way that supports further ratcheting up in complexity across historical time.

It may very well be, then, that it is these processes and aspects (cultural learning, teaching, normativity) that give human cultural traditions their extraordinary stability and cumulativity over time. Integral in all of these is a kind of social engagement depending on skills and motivations for shared intentionality.

## 16.4 Joint Attention and Perspective

It is, thus clear that human infants, before they are fully participating members of a culture, already have a special motivation for sharing experiences with other persons, and they possess special skills for creating with others joint goals, joint intentions, and joint attention. They learn from others in unique ways as well.

However, our claim goes further. Our claim is that participation in interactions involving shared intentionality transforms human cognition in fundamental ways. Most importantly, it actually creates new forms of cognitive representation, specifically, perspectival or dialogic cognitive representations (see also Tomasello 1999; Tomasello et al. 2005). In understanding and internalizing an adult's intentional states, including those directed towards her, at the same time she experiences her own intentional states towards the other, the child comes to conceptualize the interaction simultaneously from both the first and third person's perspective (see Barresi and Moore 1996) – forming a bird's eye view of the collaboration in which both commonalities and differences are all comprehended within a single representational format. Such perspectival representations are necessary not only for supporting cooperative interactions online, but also for the creation and use of certain kinds of cultural artifacts, most importantly linguistic and other kinds of symbols, which are socially constituted and bi-directional in the sense of containing simultaneously the perspective\_s\_of\_both\_speaker and listener (see Mead 1934).

These perspectival cognitive representations pave the way for later uniquely human cognitive achievements. Importantly, following Harris (1996), Tomasello and Rakoczy (2003) argued and presented evidence that coming to understand false beliefs – the fact that someone else's cognitive perspective about a state of affair is different from what I know to be true – depends on children's participation over a several year period in perspective-shifting discourse. In such linguistic discourse, including such things as misunderstandings and requests for clarification, children experience regularly that what another person thinks is often different from what they think, and the understanding of false beliefs – which, in almost everyone's account, is fundamental to mature human social cognition – is apparently unique to humans (Call and Tomasello 2008). And at age 4, children not only come to understand that others might hold false beliefs, they develop a sophisticated understanding of perspectives more generally. That is, they appreciate that different people might see or conceptualize a given event or object in different ways and also, that one and the same person can view or construe an object differently at different times. This ability to simultaneously “confront” perspectives becomes manifest in a variety of tasks besides the standard false belief task. For example, children now understand that one and the same object might (1) look like a rock but really be a sponge (the so-called appearance-reality distinction, Flavell et al. 1986), (2) be both an animal and a rabbit (as shown by their acceptance of alternative labels for a given object, Doherty and Perner 1998), and (3) be seen right-side-up from one perspective but upside-down from another (the so-called level 2 perspective-taking, Masangkay et al. 1974).

We would argue that young children come to understand and operate with the concept of perspective only after first experiencing the sharedness of attention on one and the same thing (see also Barresi and Moore 1996). From thereon, they can later begin to understand the differences in perspectives which converge on the shared target or object of interest. The foundations again can be found in infancy.

Evidence for this comes from a series of studies in which infants must determine what an adult has experienced and has not experienced. Tomasello and Haberl (2003) had 12- and 18-month-old infants play 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 *that* toy (without indicating by gazing or pointing which specific toy she was attending to). Surprisingly, infants of both ages selected the toy the adult had not experienced (the one which was new for her). In order to solve this task, infants had to understand (1) that people get excited about new, not familiar things and (2) which of the toys was new for the adult and which she was already familiar with from previous experience.

In this study, infants knew what was familiar for the adult after they had participated with her in joint attention around two of the objects (but not the third). This suggests the possibility that infants need to attend to another person's experience in joint attention with her in order to register the other as knowing the object in question. And this is what was basically found in the two studies by Moll et al. (2007), Moll and Tomasello (2007b). Following the basic procedure of Tomasello and Haberl (2003), 14- and 18-month-old infants either (1) became familiar with the first two objects in a joint attentional frame together with the adult or (2) simply witnessed the adult become familiar with the known objects individually. In each case, infants themselves became equally familiar with all three objects, as in the original study. The result was that infants knew which of the three objects was new for the adult, and thus captured her attention only when they had explored the known objects in a joint attentional format with her. They could not make this distinction when they had just witnessed her exploring them on her own, outside of a joint attentional frame. The shared attention to the known objects thus highlighted the fact that the third object was not jointly experienced. It is, thus inside of joint attentional frames, that infants first begin to realize differences in people's experiences and perceptions. This early understanding of other's experiences is the foundation for the later developing understanding of divergent perspectives on one and the same thing – an understanding which is, just like the ability to jointly engage with others, uniquely human.

Our argument is thus that basically all species-unique aspects of human cognition reflect their cooperative roots in fundamental ways. The ability to take the perspective of others – which spawns the understanding of false beliefs, perspectival cognitive representations, and collective/institutional reality – is only possible for organisms that can participate in social interactions involving shared intentionality, especially involving joint attention. Let us be very clear on this point. Participation in these interactions is critical. A child raised on a desert island would have all of the biological preparations for participation in interactions involving shared intentionality, but because the child did not actually participate in such interactions, she would have nothing to internalize into perspectival cognitive representations. Ontogeny in this case is critical.



## 16.5 From Collaboration to Culture

We, thus find that human infants in their second year of life are much more skilled, and much more motivated, than are great apes at participating in collaborative problem solving and cooperative communication, and their skills of social/cultural learning have unique qualities as well. Following Tomasello et al. (2005), our claim is that the reason for this difference is that human infants are biologically adapted for social/cultural interactions involving shared intentionality. Even at this tender age, human infants already have special skills for creating with other persons joint goals, joint intentions, and joint attention, and special motivations for helping and sharing with others – and for communicating with and learning from others within these special interactions as well.

It would seem that we are still a long way from such things as governments and religions and sciences and other large-scale cultural institutions. But actually we are not. These kind of cultural institutions arose only very recently in human evolution, after, and partially as a result of, the agricultural revolution when people began living in large cities. There, many new demands arose for collaborating with strangers and in much larger groups and across longer spans of time than previously. Our simple proposal is that the collaborative and communicative skills we see in young children form the necessary foundation for beginning to participate in such large-scale collaboration. Virtually no one believes that there were any genetic events at the agricultural revolution that led to the unprecedented population explosion and flowering of cultural institutions associated with that great event. So what is required for scaling up small group collaboration into large group collaboration is mainly certain sociological conditions, as a first step, and then cultural-historical processes over time and generations in which such institutions could be formed.

What we have called the cultural intelligence hypothesis, or the Vygotskian intelligence hypothesis, reflects the idea that the incredible complexities of modern human culture and its institutions are the result of a qualitatively new process that arose in human evolution. Although there may be some cultural transmission in some animal species, nothing like the human creation of cultural artifacts and institutions – which ratchet up in complexity over time and within which children's ontogeny proceeds and on which it depends – takes place in other species. This is because human beings collaborate with, communicate with, and learn from their groupmates based on unique skills and motivations for shared intentionality and cultural learning. We do not believe that the human gap can be explained by any appeal to individual cognitive skills, but rather it can only be explained by the social-cultural processes for which they are specially adapted and within which each generation of modern humans has evolved.



## References

- Barresi J, Moore C (1996) Intentional relations and social understanding. *Behav Brain Sci* 19:107–154
- Bates E (1979) *The emergence of symbols: cognition and communication in infancy*. Academic, New York
- Bates E, Camaioni L, Volterra V (1975) The acquisition of performatives prior to speech. *Merrill Palmer Q* 21:205–226
- Behne T, Carpenter M, Tomasello M (2005) One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Dev Sci* 8:492–499
- Boesch C (1991) Teaching among wild chimpanzees. *Anim Behav* 41:530–532
- Boesch C (2005) Joint cooperative hunting among wild chimpanzees: taking natural observations seriously. *Behav Brain Sci* 28:692
- Boesch C, Boesch H (1989) Hunting behavior of wild chimpanzees in the Tai-National-Park. *Am J Phys Anthropol* 78:547–573
- Boesch C, Boesch-Achermann H (2000) *The chimpanzees of the Tai forest: behavioural ecology and evolution*. Oxford University Press, Oxford
- Bratman ME (1992) Shared cooperative activity. *Phil Rev* 101:327–341
- Call J, Tomasello M (1994) The production and comprehension of referential pointing by orangutans. *J Comp Psychol* 108:307–317
- Call J, Tomasello M (2005) What chimpanzees know about seeing revisited: an explanation of the third kind. In: Eilan N, Hoerl C, McCormack T, Roessler R (eds) *Joint attention*. Oxford University Press, Oxford, pp 45–64
- Call J, Tomasello M (2008) Does the chimpanzee have a theory of mind? 30 years later. *Trends Cogn Sci* 12:187–192
- Camaioni L (1993) The development of intentional communication: a re-analysis. In: Nadel J, Camaioni L (eds) *New perspectives in early communicative development*. Routledge, New York, pp 82–96
- Carpenter M (2006) Instrumental, social, and shared goals and intentions in imitation. In: Rogers SJ, Williams J (eds) *Imitation and the development of the social mind: lessons from typical development and autism*. Guilford, New York, pp 48–70
- Carpenter M, Tomasello M, Striano T (2005) Role reversal imitation and language in typically developing infants and children with autism. *Infancy* 8:253–278
- Clark HH, Brennan SE (1991) Grounding in communication. In: Resnick LB, Levine JM, Teasley SD (eds) *Perspectives on socially shared cognition*. APA Books, Washington, DC, pp 127–149
- Doherty M, Perner J (1998) Metalinguistic awareness and theory of mind: just two words for the same thing? *Cogn Dev* 13:279–305
- Flavell JH, Green FL, Flavell ER (1986) Development of knowledge about the appearance-reality distinction. *Monogr Soc Res Child Dev* 51(Serial No 212):i–v 1–87
- Gergely G, Csibra G (2006) Sylvia's recipe: the role of imitation and pedagogy in the transmission of cultural knowledge. In: Enfield NJ, Levinson SC (eds) *Roots of human sociality: culture, cognition, and interaction*. Berg, Oxford, pp 229–255
- Hare B, Tomasello M (2004) Chimpanzees are more skillful in competitive than in cooperative cognitive tasks. *Anim Behav* 68:571–581
- Harris P (1996) Desires, beliefs, and language. In: Carruthers P, Smith PK (eds) *Theories of theories of mind*. Cambridge University Press, Cambridge, pp 200–222
- Herrmann E, Call J, Hernández-Lloreda MV, Hare B, Tomasello M (2007) Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science* 317:1360–1366
- Herrmann E, Call J, Hernández-Lloreda M, Hare B, Tomasello M (in press) The structure of individual differences in the cognitive abilities of children and chimpanzees. *Psychological Science*

- Itakura S, Agnetta B, Hare B, Tomasello M (1999) Chimpanzees use human and conspecific social cues to locate hidden food. *Dev Sci* 2:448–456
- Jensen K, Hare B, Call J, Tomasello M (2006) What's in it for me? Self-regard precludes altruism and spite in chimpanzees. *Proc Roy Soc Lond B* 273:1013–1021
- Kaminski J, Call J, Tomasello M (2004) Body orientation and face orientation: two factors controlling apes' begging behavior from humans. *Anim Cogn* 7:216–223
- Kelemen D (1999) Functions, goals and intentions: children's teleological reasoning about objects. *Trends Cogn Sci* 12:461–468
- Leavens DA, Hopkins WD (1998) Intentional communication by chimpanzees: a cross-sectional study of the use of referential gestures. *Dev Psychol* 34:813–822
- Leavens DA, Hopkins WD (2005) Multimodal concomitants of manual gesture by chimpanzees (*Pan troglodytes*): influence of food size and distance. *Gesture* 5:73–88
- Leavens DA, Hostetter AB, Wesley MJ, Hopkins WD (2004) Tactical use of unimodal and bimodal communication by chimpanzees, *Pan troglodytes*. *Anim Behav* 67:467–476
- Liebal K, Behne T, Carpenter M, Tomasello M (2009) Infants use shared experience to interpret pointing gestures. *Dev Sci* 12:264–271
- Liszkowski U, Carpenter M, Henning A, Striano T, Tomasello M (2004) Twelve-month-olds point to share attention and interest. *Dev Sci* 7:297–307
- Liszkowski U, Carpenter M, Striano T, Tomasello M (2006) 12- and 18-month-olds point to provide information for others. *J Cogn Dev* 7:173–187
- Masangkay ZS, McCluskey KA, McIntyre CW, Sims-Knight J, Vaughn BE, Flavell JH (1974) The early development of inferences about the visual percepts of others. *Child Dev* 45:357–366
- Mead GH (1934) *Mind, self, and society*. University of Chicago Press, Chicago, IL
- Melis AP, Hare B, Tomasello M (2006) Chimpanzees recruit the best collaborators. *Science* 311:1297–1300
- Moll H, Tomasello M (2007a) Cooperation and human cognition: the Vygotskian intelligence hypothesis. *Phil Trans Roy Soc B* 362:639–648
- Moll H, Tomasello M (2007b) How 14- and 18-month-olds know what others have experienced. *Dev Psychol* 43:309–317
- Moll H, Carpenter M, Tomasello M (2007) Fourteen-month-olds know what others experience only in joint engagement. *Dev Sci* 10:826–835
- Povinelli DJ, Nelson KE, Boysen ST (1992) Comprehension of role reversal in chimpanzees: evidence of empathy. *Anim Behav* 43:633–640
- Rakoczy H, Warneken F, Tomasello M (2008) The sources of normativity: young children's awareness of the normative structure of games. *Dev Psychol* 44:875–881
- Savage-Rumbaugh ES (1990) Language as a cause-effect communication system. *Philos Psychol* 3:55–76
- Searle JR (1995) *The construction of social reality*. Free Press, New York
- Silk JB, Brosnan SF, Vonk J, Henrich J, Povinelli DJ, Richardson AS, Lambeth SP, Mascaró J, Schapiro SJ (2005) Chimpanzees are indifferent to the welfare of unrelated group members. *Nature* 437:1357–1359
- Tomasello M (1999) *The cultural origins of human cognition*. Harvard University Press, Cambridge, MA
- Tomasello M (2003) *Constructing a language: a usage-based theory of language acquisition*. Harvard University Press, Cambridge, MA
- Tomasello M (2006) Why don't apes point? In: Enfield NJ, Levinson SC (eds) *Roots of human sociality: culture, cognition and interaction*. Berg, New York, pp 506–524
- Tomasello M (2008) *Origins of human communication*. MIT, Cambridge, MA
- Tomasello M, Carpenter M (2005) The emergence of social cognition in three young chimpanzees. *Monogr Soc Res Child Dev* 70:1–132
- Tomasello M, Haberl K (2003) Understanding attention: 12- and 18-month-olds know what is new for other persons. *Dev Psychol* 39:906–912

- Tomasello M, Rakoczy H (2003) What makes human cognition unique? From individual to shared to collective intentionality. *Mind Lang* 18:121–147
- Tomasello M, Kruger AC, Ratner HH (1993) Cultural learning. *Behav Brain Sci* 16:495–552
- Tomasello M, Call J, Gluckman A (1997a) Comprehension of novel communicative signs by apes and human children. *Child Dev* 68:1067–1080
- Tomasello M, Call J, Warren J, Frost GT, Carpenter M, Nagell K (1997b) The ontogeny of chimpanzee gestural signals: a comparison across groups and generations. *Evol Comm* 1:223–259
- Tomasello M, Carpenter M, Call J, Behne T, Moll H (2005) Understanding and sharing intentions: the ontogeny and phylogeny of cultural cognition. *Behav Brain Sci* 28:675–735
- Tuomela R (2002) Collective goals and communicative action. *J Phil Res* 27:29–64
- Uzgiris IC (1981) Two functions of initiation during infancy. *Int J Behav Dev* 4:1–12
- Vygotsky LS (1978) *Mind in society: the development of higher psychological processes*. Harvard University Press, Cambridge, MA
- Warneken F, Tomasello M (2006) Altruistic helping in human infants and young chimpanzees. *Science* 311:1301–1303
- Warneken F, Tomasello M (2007) Helping and cooperation at 14 months of age. *Infancy* 11: 271–294
- Warneken F, Chen F, Tomasello M (2006) Cooperative activities in young children and chimpanzees. *Child Dev* 77:640–663
- Whiten A. Coming of age for cultural Panthropology. In: Lonsdorf E, Ross S, Matsuzawa T (eds) *The mind of the chimpanzee*. University of Chicago Press, Chicago, IL (in press)