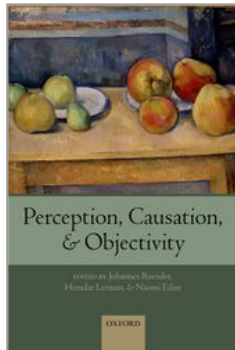


# Perspective-Taking and its Foundation in Joint Attention

---

University Press Scholarship Online

Oxford Scholarship Online



## Perception, Causation, and Objectivity

Johannes Roessler, Hemdat Lerman, and Naomi Eilan

Print publication date: 2011

Print ISBN-13: 9780199692040

Published to Oxford Scholarship Online: September 2011

DOI: 10.1093/acprof:oso/9780199692040.001.0001

## Perspective-Taking and its Foundation in Joint Attention

Henrike Moll

Andy Meltzoff

DOI:10.1093/acprof:oso/9780199692040.003.0016

### Abstract and Keywords

We propose a new developmental model that unites two phenomena that have so far been studied in isolation: joint attention in infancy and perspective-taking in young childhood. In this model, infants' abilities to jointly attend to objects and events with others, even though it does not require any understanding of perspectives, provides the necessary foundation and sets the stage for the later emerging ability to understand perspectives. Any perspectival difference presupposes a shared object onto which the perspectives converge—joint attention constitutes this shared object of perception. The understanding of perspectives then develops in two distinct steps. First, infants and young children learn to *take* perspectives, which allows them to understand others' speech acts and actions involving (perceptual, epistemic, or conceptual) perspectives that differ from their own. Second,

# Perspective-Taking and its Foundation in Joint Attention

---

children between 4 and 5 years of age come to *confront* perspectives: they can now explicitly acknowledge that the same object may be viewed or construed in alternative ways. This way of looking at children's developing social cognition sheds new light on the 'old' problem of theory of mind.<sup>t</sup>

*Keywords:* social cognition, joint attention, perspective-taking, theory of mind, cognitive development

Unlike any other species, humans can think about perspectives that are not currently their own. They can put themselves in the “mental shoes” of others and imagine how they perceive, think, or feel about an object or event. Perspectivity in its mature, adult form even goes beyond the ability to determine a specific person's point of view at a certain moment in time. It entails the general comprehension that one and the same thing or event can be viewed or construed differently depending on one's standpoint—whether this is a visuo-spatial, epistemic, conceptual, or affective standpoint (Perner, Brandl, and Garnham, 2003). From a developmental perspective, the question arises when and how children acquire this knowledge.

We offer a new look at the early ontogeny of understanding visual perception and experiences—with a major emphasis on the ability to take and understand the perspectives of others. The central claim we aim to develop is that human children first learn about perspectives within the context of joint attentional engagement. Infants' ability and motivation to jointly attend to objects and events with others allows them to share perceptions and experiences from very early on in life (Tomasello, Call, Carpenter, Behne, and Moll, 2005). This sharing sets the ground for later perspective-taking.

Developmental inquiries of joint attention and perspective-taking have mostly been conducted in separation: The term “joint visual attention” is often used as a synonym for the specific case of gaze following, which is rarely looked at in terms of its relation to later perspective-taking; and models of perspective-taking have failed to recognize early joint attentional skills as a foundational first step towards perspectivity (but see Martin, Sokol, and Elvers, 2008).

# Perspective-Taking and its Foundation in Joint Attention

---

We want to bring these two strands together and argue that joint attention is a necessary condition for appreciating perspectives. Perspectival differences, however distinct and incompatible they may be—in the sense that they cannot consistently be held by one person at the same time—necessarily converge on one and the same object (where “object” can refer to a thing, an event, a state of affairs etc.). They thus have at their (p.287) basis a common ground, and this common ground is constituted by the joint attention devoted to the object by two or more individuals.

We propose a series of social-cognitive steps taken by infants and young children on their way to a mature understanding of perspectives.<sup>1</sup> Our model overlaps in some respects with previous stage models, such as that by Flavell and colleagues (see Flavell 1978, 1992, for overviews) or Selman (1980). But it differs from these accounts in important ways. First and foremost, we acknowledge the joint attentional abilities demonstrated by infants at around one year of age as a staging post for the emergence of perspectivity. At this stage, infants are at “level 0 perspective-taking”: they do not yet know anything about perspectives, but they can share them in joint attention or joint engagement with others—as evidenced by such behaviors as gaze following, alternating gaze between object and co-attender, holding up and showing, or pointing to objects or events. This sharing of attention is qualitatively and structurally different from the tempo-spatial co-ordination of behavior that is found in primates (see Moll and Tomasello, 2007a) and it lays the grounds for the more complex forms of taking and understanding perspectives that follow during the next months and years in young childhood. About a year later, at around 2 years, children reach “level 1 visual perspective-taking”: they know *what*, e.g. which objects in a room, others can and cannot see from their current visuo-spatial viewpoint (at least when the spatial requirements, e.g. projective geometry, are minimal). We will compare this skill with infants' knowledge about what someone is or is not familiar with from past experience—which may analogously be called “level 1 experiential perspective-taking.” Counterintuitively, the latter seems to develop significantly earlier than level 1 visual perspective-taking. We think that this puzzling

developmental order may reflect something about infants' early engagement with others and the world, and about a particular challenge tied specifically to the understanding of visual perception as opposed to more holistic ways of engaging with or experiencing objects in the world. At level 2, children know what but also *how* others see things. They understand the specific way in which something is seen, construed, or (re)presented. However, in the light of new data, we argue for a division of this level in two distinct sublevels: At level 2A, which seems to be reached by 3 years, a child is able to recognize how another sees something, even when this differs from how the child sees at that moment. Yet, this does not entail the ability to “confront” perspectives and comprehend that one object, event etc. can be seen in multiple ways depending on one's viewpoint. This, as is evidenced by many theory of mind studies, seems to emerge at around 4.5 years of age, when preschoolers gain an explicit knowledge about perspectives in the various domains, including perception and knowledge. This full-blown acknowledgement of perspectives is achieved at level 2B.

## (p.288) 1 Level 0 perspective-taking: Sharing attention

As early as in the first year, human children can share their visual attention with others. This ability manifests itself in two chief ways, both of which emerge at around the same time between 9 and 12 months of age. At this age, infants begin to *follow into* another person's already established focus of attention and *direct* another person's attentional focus by pointing to or showing the person excitedly an object of their interest (see, e.g. Carpenter, Nagell, and Tomasello, 1998). Before producing linguistic utterances to direct others' attention to things verbally, infants thus know how to use a variety of non-linguistic means to achieve a “meeting of minds” with others (Bruner, 1995). We chose to focus on the cases of gaze following (as one way of participating in joint attention) and pointing (as one way of initiating joint attention) as these have been subjected to numerous experimental investigations.

## 1.1 Gaze following

Gaze following is probably the most widely investigated behavior that often marks the beginning of a joint attentional sequence. In the first half of their first year of life, human infants look in the general direction another person is looking (e.g. D'Entremont, Hains, and Muir, 1997; Scaife and Bruner, 1975). However, at this young age they only follow gaze to a target when it is inside their visual field and the first object on their scan path, suggesting the possibility that infants are simply orienting in the same direction in which another's head is oriented (Butterworth, 1983). By around 12 months, however, gaze following is flexible and robust and seems to reflect an understanding that people do not just look in some direction, but rather, that they see things where they look. This may also be evidenced by the fact that infants' gaze following is not limited to their immediate visual fields: they locomote behind barriers to see what others have just seen there (Moll and Tomasello, 2004) and follow gaze to the space behind their own bodies (Deák, Flom, and Pick, 2000). Other clever variations of the classic gaze following have established that shortly after their first birthdays, infants have implicit knowledge that the eyes play a critical role in seeing, i.e. they need to be oriented towards the object, open (Brooks and Meltzoff, 2002; Meltzoff and Brooks, 2008), and that opaque barriers on the visual scan path to objects prevent people from seeing these objects (Caron, Kiel, Dayton, Butler, 2002).

However, humans are not the only species that align their regard with that of a conspecific or human. Not only non-human primates follow gaze to where another has just looked (even behind barriers, Tomasello, Hare, and Agnetta, 1999), but also dolphins (Pack and Herman, 2004), goats (Kaminski, Riedel, Call, and Tomasello, 2005), and ravens (Schloegl, Kotrschal, and Bugnyar, 2007). The mere behavior of “looking where someone else is looking” (Butterworth, 1991, p. 223) then does not necessarily indicate a sharing of attention or experiences—unless one wanted to attribute joint attention to all these species. We think it is important to broaden the (p.289) scope and take into account the social context in which gaze following behavior occurs in human infants, its distinct phenomenal quality, and the alternative ways in which

# Perspective-Taking and its Foundation in Joint Attention

---

infants not only participate in, but also actively establish joint attention by gesturally inviting others to share experiences with them.

Having followed an adult's gaze, infants frequently point to the object, vocalize (Brooks and Meltzoff, 2002), or look back to the adult (Carpenter et al., 1998). These “checking looks” to the other person close the circuit of the triangulation by providing the arrow that connects the two co-attenders with each other. Additionally, infants often show a “knowing smile” as they look to the other's face—thereby demonstrating an awareness of the mutuality of the experience. More so, this may be a manifestation of the infants' feeling of “interpersonal connectedness” and identification with the other. It certainly means that there is not just an identical target attended to separately by two individuals at the same time, but that we have here an instance of two people sharing an orientation or attitude towards an object (Hobson, 2005). Furthermore, gaze following in humans usually initiates or continues an extended joint attentional episode or “format” (Bruner, 1983). These are the mundane, simple cooperative activities shared by caregiver and infant such as sharing a meal, taking a bath, playing, engaging in simple problem-solving tasks (see Heal, 2005). A lot of these activities have a structure of reciprocal role- and turn-taking (give-take, hide-seek, etc.). Roles and perspectives can be thought of as equivalent constituents in the action and the perceptual domain: just as role-taking is learned within the context of simple cooperative activities, so is perspective-taking acquired within joint attentional sequences. In any case, gaze following is just a snapshot extracted from a longer scene in which infant and adult alternate gaze, vocally “comment” on the object and engage in shared experiences and explorations.

## 1.2 Pointing

At around the same time as infants follow into others' attentional focus, they also direct others' attention by pointing. While non-human primates seem to point only imperatively, using others as “social tools” (Bates, Camaioni, and Volterra, 1975) to get them what they want (Tomasello, 2006; but see Racine, Leavens, Susswein, and Wehera, 2008),

# Perspective-Taking and its Foundation in Joint Attention

---

human infants point for a variety of motives. They often point not to request things, but simply to share their experience of something with another person—what has been called declarative pointing (see Bates et al., 1975). This may be most obvious when “requesting” cannot be what the child attempts to do, for example because what she points to is i) well within her own reach, so that no help retrieving it would be required, or ii) far outside of hers as well as the adult's reach (e.g. a plane in the sky), such that no such help can be expected, or iii) not a thing, but an event or a state of the world.

One may still object that what looks like a sharing motive at first is really an imperative one—for example the desire to attract another's attention, to get the other to attend to oneself (Moore and Corkum, 1994). However, experimental data suggest that this is unlikely. In one experiment, pointing gestures were elicited (p.290) in 12-month-old infants by showing them an event an adult did not attend to at first. The infants were satisfied and ceased to point for the adult only when the adult alternated gaze and truly shared the interesting sight with them. By contrast, if the adult either attended to the infant or the event alone, infants were dissatisfied and persisted to point (Liszkowski, Carpenter, Henning, Striano, and Tomasello, 2004). Thus, infants urged the other to close the triangulation and share her orientation to the attended-to event. Just witnessing the other establish a “parallel” instead of a joint engagement with the event alone, or, witnessing the other establish *dyadic* engagement with them alone was not satisfactory. Other studies have shown that infants complement their gestures with looks to the adult's face to check if their point is received and acknowledged—with these looks changing from being “reactive” to being anticipatory over the course of the first half of the second year of life (Franco and Butterworth, 1996). It seems that a whole variety of motives to point for others are “buried” under the label of declarative pointing, so that the category is perhaps better seen as “non-imperative pointing,” where non-imperative motives include: providing others with information about the presence/status of something, sharing simply for the sake of sharing, requesting information such as the name or function of the pointed-to object etc. (see also

Tomasello, Carpenter, and Liszkowski, 2007; Tomasello, 2008).

## 1.3 Synopsis

By around one year of age, human infants establish joint visual attention by both tuning into another's pre-established focus of attention as well as inviting others to share theirs. Some of the behaviors we would call joint visual attention in human children, such as gaze and point following, are also found in non-human animals, but the social scenarios in which they are embedded, their phenomenal quality, their bi-directionality (the ability to take the role of the initiator and follower in joint attention and switch between them) and the diversity of motives are clearly distinctive in human-human-interactions. For these reasons, we feel confident to say that the infant in these situations is aware that the other shares her attentional focus—a criterion for joint attention which most philosophers and psychologists seem to agree upon (see Eilan, Hoerl, McCormack, and Roessler, 2005). However, we do not think that at this early point the sharing of attention reflects an understanding of perspectives or perspectival differences. As Barresi and Moore (1993) have put it the “sharing of perspectives precedes the understanding of these perspectives” (p. 513). In this regard, we follow a philosophical tradition that construes the early joint attentional abilities of infants as a form of “knowing how” rather than “knowing that” (e.g. Seemann, 2007). Participating and engaging in joint attention is primarily an “empractical” (Bühler, 1965; Stekeler-Weithofer, 2005) skill. The use of this skill then allows for and blossoms into the development of the more complex forms of perspective-taking and understanding gained in the next months and years of life.

## (p.291) 2 Level 1 Perspective-taking

In level 1 perspective-taking according to Flavell (e.g. 1992) and colleagues' framework, a child not only recognizes others' attention, but also knows *what* others can and cannot visually perceive in the moment (visual perspective-taking). In other words, a child at this level knows what objects do and do not figure in another's visual perspective. There is also an analogous level of understanding that has received much



# Perspective-Taking and its Foundation in Joint Attention

---

attention in developmental research in the past years, namely the understanding of what others have and have not become familiar with from past perceptual experience. As will become clear in the following sections, these two abilities are quite distinct and challenge infants and young children to different degrees.

## 2.1 Level 1 visual perspective-taking

According to Flavell and colleagues' framework, level 1 visual perspective-taking starts with an understanding of what others can (and cannot) see from their specific viewpoint. A child who has reached "level 1 visual perspective-taking" should be able to know what objects can and cannot be seen from a certain visuo-spatial position. The gaze following procedure and its variations are informative about infants' implicit knowledge about some basic enabling and defeating conditions of seeing (e.g. that the eyes need to be open and the line of sight clear)—but are not appropriate measures when it comes to determining if a child knows what is and is not part of another person's perspective. Richer response measures are required for this. The child needs to specify somehow—verbally, gesturally, or by complying to a request with some sort of action—exactly what can or cannot be seen from a certain spatial position.

A seminal study was conducted by Masangkay, McCluskey, McIntyre, Sims-Knight, Vaughn, and Flavell (1974). In their experiment, an adult held up a card between herself and the child. The side of the card facing the child contained a picture of one animal, e.g. a dog, while the side facing the adult showed a different animal, e.g. a cat. The child was previously shown both sides of the card and so knew what each side depicted. She was then asked what she herself saw and what the adult saw. Most children at the age of 2.5 years and older could say correctly what they saw and what the adult saw.

Two other studies have provided converging evidence that level 1 perspective-taking develops at around 2.5 years—but also point at some limitations at this age. In a study by Flavell, Shipstead, and Croft (1978) a child was asked to hide an object from an adult (who sat either next to or across from the child) by either placing an object in relation to a barrier that

was already on the table or by placing a barrier in relation to an object that was already on the table. The youngest age group of 2.5 year-olds successfully placed the toy on the table so that it was hidden from the adult's view, but not from themselves. However, only children 3 years and older knew how to “interrupt” an adult's already established visual engagement with an object by positioning a barrier between the object and the adult (see also McGuigan and Doherty, 2002).

(p.292) In a study using a search paradigm, Moll and Tomasello (2006) found that 24-month-olds have a nascent understanding of what others can and cannot see from their viewpoint. An adult pretended to be searching for an object. There were two candidate objects in the room, both of which were well visible and equidistant from the child position. Behind (from the child's perspective) one of the objects was an opaque barrier which blocked the adult's view to it. The 24-month-olds selected this object significantly in response to the adult's searching, but had no preference for this object in a control condition in which the adult made a neutral and ambiguous request for an object. The children thus knew i) that people search for things they cannot see, and ii) which of the two objects in this situation could not be seen by the adult.

The research suggests that young children begin to appreciate that others may not see what they see at around 2 to 2.5 years of age.<sup>2</sup> It is not surprising that this ability comes into place significantly later than gaze following and other level 0 skills. More surprising, however, is the finding that level 1 visual perspective-taking is *preceded*, not succeeded, as will be shown in the following section, by the understanding of what others are and are not familiar with from past experience.

## 2.2 Level 1 experiential perspective-taking

Recent research suggests that children can understand what others know and do not know at a surprisingly young age; at least if “know” refers not to propositional knowledge, but to the type of knowledge that is conveyed by “connaître” in French, “kennen” in German, and “conocer” in Spanish—which is probably best translated with “being familiar” or “acquainted” with something from past perceptual experience.

# Perspective-Taking and its Foundation in Joint Attention

---

In O'Neill's (1996) well-known study, a child saw an experimenter place a desirable object in one of two containers out of the child's reach. The parent, who had the role of the child's helper, either witnessed the hiding event or missed it because she was out of the room or covered up her eyes (in which case the child was explicitly told that she cannot see). Children of 2.3 and 2.7 years of age tailored their requests for the parent according to her knowledge state. If the parent was ignorant, they made more frequent and more specific requests than when she knew where the object was placed.

A series of recent studies shows that even infants early in the second year of life can judge what others are and are not familiar with. In a study by Tomasello and Haberl (2003), 12- and 18-month-old infants and an adult jointly engaged with two novel objects in turn for one minute each. Then the adult left the room. While she was gone, the infant and a second adult played with a third novel object. Finally, all three objects were held in front of the infant, at which point the first adult returned and excitedly exclaimed "Wow! Look! Look at that one!" gazing in the direction of all three objects. She then made an ambiguous request for the infant to hand "it" to her. Both the (p.293) 12- and 18-month-olds significantly chose the third object—but not in a control condition in which the adult experienced all three objects. The infants thus knew what the adult did and did not experience, independently from their own experience.

Importantly, this understanding is not limited to the specific novelty paradigm. In other tasks, infants of 14 months and older were equally able to i) select an object that was mutually familiar, but had been shared in special ways between infant and adult prior to her making an ambiguous request for "it" (Moll, Richter, Carpenter, and Tomasello, 2008), and ii) see an adult's expression of excitement as being directed at either an entire object or a part of the object, depending on whether the adult saw it for the first time or knew it from prior experience (Moll, Koring, Carpenter, and Tomasello, 2006).

## 2.3 Making sense of a puzzling developmental order

The procedures that have been developed to assess level 1 visual and experiential perspective-taking seem highly similar:

In addition to one or more objects that are mutually seen/known, there is another object which only the child sees/knows but the adult does not. In both cases the question is if the child can ignore her perception of/familiarity with that object and recognize the adult's ignorance of it. Yet, children are successful in the 'experiential task' almost a year before they solve the visual perception task. Intuitively, one would assume that children come to know what others can and cannot see "in the here and now" *before* they come to know what others do and do not know from previous experience—which involves keeping track of what happened in the recent past. The pressing question then is why the recognition of past experiences develops so early, and why level 1 visual perspective-taking emerges relatively late.

### **2.3.1 Social engagement facilitates recognizing what others experience**

In an attempt to reconcile these seemingly contradictory findings, Moll and Tomasello (2007b) put forth the "sharing hypothesis". They argued that what enabled infants to perform well in knowledge-ignorance studies was the "sharing" of the two known objects: being jointly engaged with the adult as she explored the familiar objects allowed them to register the adult as knowing the objects a few moments later. The unknown object stuck out as the one that the infant and adult had not shared together. To test this hypothesis, Moll and Tomasello (2007b) varied the specific way in which the adult became familiar with the two known objects. In one condition—modelled on Tomasello and Haberl's experimental condition—the adult shared her experience of the two known objects with the infant in joint engagement. In two other conditions, (1) infants observed the adult examine the two known objects individually instead of in joint engagement, or (2) the adult looked on from afar as the infant and the assistant examined the two familiar objects. As in Tomasello and Haberl's (2003) study, the adult then left the room while the assistant presented the infant with the third object.

In line with the hypothesis, 14-month-old infants knew which object was new for the adult only when they had shared the experience of the known objects together. In (p.294) both other conditions in which the objects were not shared, infants

# Perspective-Taking and its Foundation in Joint Attention

---

failed to identify what the adult was referring to in her excited request. (By 18 months, infants knew what the adult had experienced not just through joint attentional engagement, but also by observing the adult actively manipulate the known objects.)

More empirical support for the view that infants come to understand what others experience through joint engagement stems from a study by Moll, Carpenter, and Tomasello (2007). They found that 14-month-olds failed on the test if they simply witnessed an adult jointly engaging with the familiar objects with *another* person from a third-person perspective. Instead, infants had to share the objects with the adult *directly* in order to register her as knowing them. Thus, joint engagement is at least helpful, probably even necessary for infants at 14 months to register others as having experienced objects. This is in accord with a point made by Heal (2005) about the critical importance of the second person. Children do not learn about the social world mostly and usually from third persons—“he”s and “she”s that are distantly observed from the outside. Instead, they learn from the “you”s with whom they interact and engage in collaborative activities with joint goals and shared attention. As Heal writes “the basic subjects of psychological predicates will be ‘us’: viz. you and me” (2005, p. 41). Only later do children learn from third parties by observing, eavesdropping, and overhearing. For example, 18-month-olds regulate their imitation of actions on an object through observing an emotional interaction between two *other* people (Repacholi and Meltzoff, 2007; Repacholi, Meltzoff, and Olsen, 2008). Likewise, infants 18 months and older learn novel words by overhearing what third persons say to each other (Floor and Akhtar, 2006). But at the beginning—and this may only be a few months prior—learning takes place strictly within the “I-thou” (Buber, 1958) relationship.

## **2.2.2 Social engagement compromises a recognition of what others do not experience**

These studies thus help to understand why or under what social conditions infants attribute experiences to others at a surprisingly young age—but they do not address the particular challenge posed by visual perspective-taking. To account for this as well, Moll, Carpenter, and Tomasello (2011) extended

# Perspective-Taking and its Foundation in Joint Attention

---

the “sharing hypothesis” and postulated that just as social engagement facilitates children's ability to recognize others' experiences, it might lead them to overestimate what has been shared, that is, it might hinder their ability to detect ignorance in others.

When a young child is engaged with another person, she might act on the presumption that she and the other person perceptually share the space around them—even though the other person cannot see what the child sees. Even adults can be “tricked” into falsely assuming a shared perceptual space with others in social situations (see, e.g. Epley, Morewedge, and Keysar, 2004). For example, a speaker might point to his computer screen instead of the projection on the wall behind him to show a graph to his audience—not realizing in that moment that the audience cannot see the laptop screen. The joint presence and social interaction suggests a shared perceptual access to the things in the room. It is possible then that children attribute ignorance to another person readily as long as the (p.295) person is not socially engaged with them at all, which is the case in classic knowledge-ignorance tasks. In these tasks, the adult disengages entirely from the situation by leaving (e.g. O'Neill, 1996; Tomasello and Haberl, 2003) or at least turning away (Southgate, Senju, and Csibra, 2007). In contrast, in visual perspective-taking tasks the adult is necessarily physically co-present. What is experimentally manipulated is not the other's presence, but merely her visual access to the objects. In such a situation it should be much harder to detect ignorance or perceptual non-connectedness, as children would need to realize that *despite* the other's co-presence, a mutual perceptual access to the objects cannot be taken for granted.

An adult's physical co-presence in the child's visual field (close by and facing the child) may be the most obvious basis for an assumption of shared experience—especially at a very young age when the objects of joint attention are mostly physical objects in the near environment. But errors of over-attributing perception and knowledge can also occur when the other is absent, but jointly engaged via verbal communication. For instance, people sometimes provide visual gestures for others

# Perspective-Taking and its Foundation in Joint Attention

---

with whom they are talking on the phone, but who cannot see them. Two-year-olds, who have just begun to be language users, may also overestimate another person's perceptual access in communicative situations.

To investigate the separate and combined effects of physical co-presence and verbal communication on children's detection of ignorance, Moll et al. (2010) again modified Tomasello and Haberl's (2003) selection paradigm. In each of four conditions, 24-month-old children shared two novel objects in turn with an adult in joint engagement, making those objects mutually known. Then, in all conditions, the third object (the target) was presented to children, but the adult never saw it. What was varied across conditions was the social situation in which children experienced the target: the adult was physically co-present or not and/or communicated verbally with children or not, in a 2x2 design. The question was if children were able to register the adult as being ignorant of the third object in these different situations. In line with the extended "sharing hypothesis," the two-year-olds over-attributed experience to the adult in all these three cases: when the adult was co-present (irrespective of whether she additionally communicated or not) and when she was absent, but communicated. Only when the adult terminated the social interaction entirely by leaving and stopping to communicate did the 2-year-olds clearly register her ignorance of the object.

What this study shows is that young children's social engagement with others may sometimes lead them to overestimate what is shared. When interacting with a co-present or communicating person, they tend to erroneously assume a shared perceptual space with that person. Just as the "curse of knowledge" compromises the ability to reason about others' false beliefs (Birch, and Bloom, 2007), so can the "curse of social engagement" compromise the ability to register others' ignorance. Importantly, this study helps to explain the discrepancy between an understanding of seeing in the here and now (visual perspective-taking) on the one hand and having experienced things in the recent past (experiential perspective-taking) on the other. While visual (p.296) perspective-taking inherently involves a co-present adult, the person in experiential perspective-taking tasks

usually breaks the social engagement entirely, e.g. by saying goodbye and leaving. In this situation, it is much easier to register others' ignorance.

From a broader perspective, this shows that what is primary is the sharedness of the situation, the “being-in-this-together.” It is likely that infants at this stage conceive of perception more holistically as someone's engagement with things. It seems that “seeing” is understood as “being engaged” or “occupied” with something (see also McGuigan and Doherty, 2002, O'Neill, 1996). Only later does the concept of seeing become refined and identified as the specific form of *visual* experience that it is, with an *explicit* understanding of the conditions for informational access and its relation to knowledge (see Wimmer, Hogrefe, and Perner, 1988). Starting at around two years, children can be brought to understand that a person may not see something despite being present and even posturally and visually oriented towards an object. But to achieve this, the other person must make very clear that there is something she cannot visually get in “contact” with—by either verbally saying that she cannot see something (as in O'Neill's, 1996, study) or by searching (as in Moll and Tomasello's, 2006, study). The impeding effect of one's co-presence can thus be counteracted by providing specific cues to one's inability to see.

### 3 Level 2 perspective-taking: Understanding “seeing as”

When a child comes to understand not only *what* is visible from a certain point of view, but also *how* a given object is seen or presented, she is considered to have reached “level 2 perspective-taking.” In philosophical terms, the child can now specify an object's *mode of presentation* or *aspectual shape* (Perner et al., 2003; Searle, 1992). For instance, an object can only be said to be “left”/“right” or “in front of”/“behind” another object as a function of one's visuo-spatial perspective. Perner and colleagues (Perner et al., 2003) have pointed out that this is the first level that strictly deserves to be called an understanding of perspectives: If a perspective is a way of



# Perspective-Taking and its Foundation in Joint Attention

---

seeing, then an understanding of perspectives necessarily entails knowledge of how people see what they see.

The most well-known level 2 perspective taking task is probably the three-mountain problem designed by Piaget and Inhelder (1956). Children sat in front of a three-dimensional model showing three mountains each with a distinctive landmark (a church etc.). A doll was placed at various positions facing the model and the child had to determine the doll's visual perspective, for example, by choosing from among a set of pictures depicting the model as seen from different viewpoints. For reasons such as the complexity of the visual array (see Borke, 1975), the use of this task has led to significant underestimations of children's capacities to imagine how an object looks from a viewpoint other than their own.

(p.297) A more child-friendly task for preschoolers is that developed by Masangkay et al. (1974). They presented children a picture of a turtle placed on the table in front of them. The children correctly identified the turtle as “right-side up” when the turtle's feet were facing them, and as “upside down” when the picture was turned so that the turtle's feet were facing away from them. However, children below 4.5 years of age did not understand that while they saw the turtle right-side up, an adult sitting across the table saw it upside-down. Replacing the word pair “upside down” and “right-side up” with the potentially more child-friendly expressions “standing on its feet” and “lying on its back” failed to improve 3-year-olds performance (Flavell, Everett, Croft, and Flavell, 1981). Other studies have looked at children's understanding of how an observer's distance from an object affects its perceived clarity and size (Flavell, Flavell, Green, and Wilcox, 1980; Pillow and Flavell, 1986). Taken together, level 2 research has consistently shown that 4.5-year-olds are mostly successful in judging how an object looks from perspectives other than their own, whereas 3-year-olds are not.

This is in line with the idea of a strong ontogenetic tie among the classic theory-of-mind abilities. Reasoning about beliefs (epistemic perspective-taking), distinguishing between appearance and reality (conceptual perspective-taking)

# Perspective-Taking and its Foundation in Joint Attention

---

accepting alternative names for a given object (e.g. “bunny” and “rabbit” for the same animal, see below), and level 2 visual perspective-taking all emerge in synchrony. They co-emerge not per coincidence, but because of conceptual relatedness: they all require an understanding that one and the same object or event can be looked at, conceptualized or interpreted in multiple ways depending on one's point of view (Perner, 2000; Perner, Stummer, Sprung, and Doherty, 2002).

Some recent studies *prima facie* seem to challenge this “unitary view,” including a series of experiments coming from our laboratory. We re-examined the development of level 2 visual perspective-taking using a color filter technique (Moll and Meltzoff, in press). An advantage of this approach may be that children at this age know the basic color terms, whereas perspectival word-pairs such as “left/right,” “in front of/behind” are not yet well understood by children this young (Wanska, 1984). In one experiment, 36-month-old children were presented with an ambiguous verbal request for an object and had to take an adult's visual perspective in order to disambiguate it. There were two candidate objects both of which the children saw in their true, same color: either white (Color Task) or blue (Color Mix Task). However, an adult saw one of them through a tinted filter—resulting in a perception of a different color for this object. Despite the fact that the children themselves saw two identically-colored objects, they systematically chose the object that the adult requested. For example, in the Color Mix Task, when the adult requested a green object, the children chose that one of the two blue objects that looked green to the adult. Moreover, children succeeded in the opposite case: they correctly chose that one of the two blue objects that the adult saw as blue (through the clear p.298 side of the screen, when the adult requested “the blue one” (Moll and Meltzoff, 2011).<sup>3</sup>

In a second experiment, children of the same age could also take an adult's perspective in a production version. The children sat next to the adult (90 degrees to her left or right) who faced a screen containing a yellow filter. The children were then requested to make a blue object look green for the adult by placing it on either side of the filter. In this production task, 36-month-old children correctly placed the

# Perspective-Taking and its Foundation in Joint Attention

---

object relative to the screen such that the adult saw it green—even though the children still saw the object in its true, blue color. This result together with the previous one again suggests that 36-month-olds understand how another person sees something when this differs from how they themselves see it.

A pressing question then is why the children in the present studies performed so well. One possibility is that 3-year-olds' understanding of visual perspectives has previously been underestimated due to extraneous task demands, such as the verbal ability to use perspectival word pairs. The new task may simply be a more sensitive measure for the same competence tested with the classic tasks. On this view, level 2 visual perspective-taking has been brought down by about 1.5 years, to 36 months of age.

This would have profound theoretical implications. Most importantly, it would undermine the idea of a common cognitive denominator shared by perceptual, conceptual, epistemic perspective-taking and so forth. One theoretical response might be to draw a distinction between different kinds of mental states such as perception and belief. Maybe visual perspectives are understood prior to epistemic ones and the challenge of classic theory-of-mind is limited to belief reasoning. In line with this view, many have argued that perception and desire are grasped by children well before epistemic states (e.g. Astington and Gopnik, 1991; Rakoczy, Warneken, and Tomasello, 2007). However, perceiving, along with believing and knowing, is considered a “cognitive attitude” with a mind-to-world direction of fit and is thus in this regard more similar to these mental states than to desires and other “conative attitudes” (which have a “world-to-mind direction of fit”, see, e.g. Gopnik, Slaughter, Meltzoff, 1994).

### 3.1 Level 2: Taking (2A) versus confronting (2B) perspectives

We would like to take a route that accommodates our findings with the unitary view. Our study may not capture perspectives at the same level that is required for an understanding of false belief, the distinction between appearance and reality, alternative (p.299) naming, and level 2 visual perspective-taking as measured by the turtle task. In these tasks children

# Perspective-Taking and its Foundation in Joint Attention

---

have to simultaneously “confront”, to borrow Perner et al.'s (2002) term, two different perspectives on the same thing. In the false belief task, the child needs to understand that another's false epistemic perspective (on an object's location or the content of a box) clashes with what she herself knows to be true. In appearance–reality tasks, two conceptual perspectives have to be confronted: the self-same object can be construed as, e.g. a rock from the “phenomenological perspective” and as a sponge from the “reality perspective.” Similarly in the alternative naming task, it needs to be acknowledged that one and the same object, e.g. a rabbit, can be conceptualized and labeled both as a “rabbit” and as a “bunny” (Doherty and Perner, 1998). Likewise in Masangkay et al.'s (1974) turtle task children have to understand that the turtle looks “upside-down” from one visual perspective but “right side-up” from another. In other words, what is put to a test is the understanding that there can be two different judgments, construals, or (re)presentations of one and the same thing held by two people at the same time.

Such a simultaneous confrontation of perspectives, however, is not necessary in the color filter tasks (Moll and Meltzoff, 2011). To succeed in these tasks, the child needs to recognize how the adult sees an object but not how that compares to their own perception of it. They can ignore the fact that what looks, for example, green to the adult looks blue to them—because they are not asked to contrast or confront the others' perspective with their own at that time.

The difference is the following: children as young as 36 months can take another's visual perspective of something even when the visual input of the same object is different for the child at that moment. In this sense, 3-year-olds engage in a form of perspective-taking that fulfills the classic definition of level 2 (e.g., Masangkay et al., 1974). However, level 2 perspective-taking has also been described as the understanding that two people may “have different perspectives or views of the same display” (Flavell, 1992, p. 119) or that an object can be seen in multiple ways. It has been taken for granted that this knowledge comes for free once a child engages in perspective-taking—the clash with the child's own perspective was simply presupposed as being

# Perspective-Taking and its Foundation in Joint Attention

---

registered by the child. But it seems that the 3-year-olds can just ignore the fact that they see the object differently from the way the adult sees it. The ability to register and reflect on perspectival differences must be seen as a distinct capacity. Two things that have been subsumed under “level 2 visual perspective-taking” thus need to be differentiated in two sublevels: the ability to take another's perspective on an object (2A) and the ability to confront two perspectives on the same object (2B). While 2A is well in place by 3 years of age, 2B emerges at around 4.5 years, as has been established and replicated in numerous false belief and similar theory of mind tasks (see e.g., Wellman, Cross, and Watson, 2001). The child has come to understand that people's relations to objects are perspectival—they understand their own perspective *as* their own and that of another *as* that of the other. They know that an object can be viewed as one thing or another—for example as a sponge or a rock in the appearance–reality test or as a duck or a rabbit (p.300) in the famous duck–rabbit figure (see Doherty and Wimmer, 2005). What our color filter tasks have surprisingly shown, *ex negativo*, is that to capture this full-blown understanding of perspectives experimentally, children have to confront two perspectives at the same time.

## Summary

Past stage models of perspective-taking have started with the ability of children to put themselves in perspectives that are different from their own. The ability of infants to share perspectives in joint attention was not seen as relevant in these accounts. In the present chapter, we argued that this early sharing of experiences in joint attention (level 0 perspective-taking) needs to be acknowledged as a staging post in the development of perspectivity, as it permits the later emergence of taking and understanding perspectives. When first taking others' perspectives, young children surprisingly find it easier to grasp what another has and has not experienced in the recent past (level 1 experiential perspective-taking) than to judge what another can and cannot see here and now (level 1 visual perspective-taking). To explain this counterintuitive order, we have again drawn on social engagement: while social engagement helps young

# Perspective-Taking and its Foundation in Joint Attention

---

children to register others' experiences with things, it sometimes leads them to overestimate the shared perceptual space and thus hinders their recognition of others' ignorance. Finally, we have urged for a distinction between two separate abilities that have so far been subsumed under level 2 (Flavell, 1992): the ability to *take* another's perspective that differs from one's own view of an object and the ability to *confront* perspectives with another (be these actual perspectives held by concrete individuals or possible perspectives that "one" could hold). New data suggest that 3-year-olds have no problems taking another's perspective (and leaving behind their own), but they still lack the ability to confront perspectives and understand the perspectival nature of people's construals of objects and events.

## References

### Bibliography references:

- Astington, J. W., and Gopnik, A. (1991). Theoretical explanations of children's understanding of the mind. *British Journal of Developmental Psychology*, 9, 7-31.
- Barresi, J., and Moore, C. (1993). Sharing a perspective precedes the understanding of that perspective. *Behavioral and Brain Sciences*, 16, 513-514.
- Bates, E., Camaioni, L., and Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly*, 21(3), 205-226.
- Birch, S. A. J., and Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science*, 18(5), 382-386.
- Borke, H. (1975). Piaget's mountains revisited: Changes in the egocentric landscape. *Developmental Psychology*, 11, 240-243.
- Brooks, R., and Meltzoff, A. N. (2002). The importance of eyes: How infants interpret adult looking behavior. *Developmental Psychology*, 38(6), 958-966.

# Perspective-Taking and its Foundation in Joint Attention

---

Bruner, J. (1983). *Child's talk: Learning to use language*. New York: Norton.

Bruner, J. (1995). From joint attention to the meeting of minds: An introduction. In C. Moore and P. J. Dunham (eds.), *Joint attention: Its origins and role in development* (pp. 1-14). Hillsdale, NJ: Erlbaum.

Buber, M. (1958). *I and thou* (R. G. Smith, trans.). New York: Charles Scribner's Sons.

Bühler, K. (1934): *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Jena: Fischer.

Butterworth, G. (1983). Structure of the mind in human infancy. *Advances in Infancy Research*, (Vol. 2, pp. 1-29). Norwood, NJ: Ablex.

Butterworth, G. (1991). The ontogeny and phylogeny of joint visual attention. In A. Whiten (ed.), *Natural theories of mind: Evolution, development, and simulation of everyday mindreading* (pp. 223-232). Oxford: Blackwell.

Caron, A. J., Kiel, E. J., Dayton, M., and Butler, S. C. (2002). Comprehension of the referential intent of looking and pointing between 12 and 15 months. *Journal of Cognition and Development*, 3(4), 445-464.

Carpenter, M., Nagell, K., and Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63 (4, Serial No. 255).

D'Entremont, B., Hains, S. M. J., and Muir, D. W. (1997). A demonstration of gaze following in 3- to 6-month-olds. *Infant Behavior and Development*, 20(4), 569-572.

Deák, G. O., Flom, R. A., and Pick, A. D. (2000). Effects of gesture and target on 12- and 18-month-olds' joint visual attention to objects in front of or behind them. *Developmental Psychology*, 36(4), 511-523.

# Perspective-Taking and its Foundation in Joint Attention

---

Doherty, M., and Perner, J. (1998). Metalinguistic awareness and theory of mind: Just two words for the same thing? *Cognitive Development*, 13, 279-305.

Doherty, M. J., and Wimmer, M. C. (2005). Children's understanding of ambiguous figures: Which cognitive developments are necessary to experience reversal? *Cognitive Development*, 20, 407-421.

Eilan, N., Hoerl, C., McCormack, T., and Roessler, J. (eds.) (2005). *Joint attention: Communication and other minds*. Oxford: Oxford University Press.

Epley, N., Morewedge, C. K., and Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, 40, 760-768.

Flavell, J. H. (1978). The development of knowledge about visual perception. In C. B. Keasey (ed.), *The Nebraska Symposium on Motivation: Vol. 25. Social cognitive development* (pp. 43-76). Lincoln: University of Nebraska Press.

Flavell, J. H. (1992). Perspectives on perspective taking. In H. Beilin and P. B. Pufall (eds.) *The Jean Piaget symposium series. Vol. 14: Piaget's theory: prospects and possibilities* (pp. 107-139). Hillsdale, NJ: Erlbaum.

Flavell, J. H., Everett, B. A., Croft, K., and Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the level 1-level 2 distinction. *Developmental Psychology*, 17, 99-103.

Flavell, J. H., Flavell, E. R., Green, F. L., and Wilcox, S. A. (1980). Young children's knowledge about visual perception. Effect of observer's distance from target on perceptual clarity of target. *Developmental Psychology*, 16, 10-12.

Flavell, J. H., Shipstead, S. G., and Croft, K. (1978). Young children's knowledge about visual perception: Hiding objects from others. *Child Development*, 49, 1208-1211.



# Perspective-Taking and its Foundation in Joint Attention

---

Floor, P., and Akhtar, N. (2006). Can 18-month-old infants learn words by listening in on conversations? *Infancy*, 9, 327-339.

Franco, F., and Butterworth, G. (1996). Pointing and social awareness: Declaring and requesting in the second year. *Journal of Child Language*, 23, 307-336.

Gopnik, A., Slaughter, V., and Meltzoff, A. N. (1994). Changing your views: How understanding visual perception can lead to a new theory of the mind. In C. Lewis, and P. Mitchell (eds.), *Children's early understanding of mind: Origins and development* (pp. 157-181). Hillsdale, NJ: Erlbaum.

Heal, J. (2005). Joint attention and understanding the mind. In N. Eilan, C. Hoerl, T. McCormack, and J. Roessler (eds.), *Joint attention: Communication and other minds* (34-44). Oxford: Oxford University Press.

Hobson, P. (2005). What puts the jointness into joint attention? In N. Eilan, C. Hoerl, T. McCormack, and J. Roessler (eds.), *Joint attention: Communication and other minds* (185-204). Oxford: Oxford University Press.

Kaminski, J., Riedel, J., Call, J., and Tomasello, M. (2005). Domestic goats, *Capra hircus*, follow gaze direction and use social cues in an object choice task. *Animal Behaviour*, 69(1), 11-18.

Keysar, B., Lin, S., and Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, 89, 25-41.

Liszkowski, U., Carpenter, M., Henning, A., Striano, T., and Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science*, 7(3), 297-307.

Luo, Y., and Baillargeon, R. (2007). Do 12.5-month-old infants consider what objects others can see when interpreting their actions? *Cognition*, 105, 489-512.

Martin, J., Sokol, B. W., and Elfers, T. (2008). Taking and coordinating perspectives: From prereflective interactivity,

# Perspective-Taking and its Foundation in Joint Attention

---

through reflective intersubjectivity, to metareflective sociality. *Human Development*, 51, 294–317.

Masangkay, Z. S., McCluskey, K. A., McIntyre, C. W., Sims-Knight, J., Vaughn, B. E., and Flavell, J. H. (1974). The early development of inferences about the visual percepts of others. *Child Development*, 45, 357–366.

McGuigan, N., and Doherty, M. J. (2002). The relation between hiding skill and judgment of eye direction in preschool children. *Developmental Psychology*, 38(3), 418–427.

Meltzoff, A. N., and Brooks, R. (2008). Self-experience as a mechanism for learning about others: a training study in social cognition. *Developmental Psychology*, 44(5), 1257–1265.

Moll, H., Carpenter, M., and Tomasello, M. (2007). Fourteen-month-olds know what others experience only in joint engagement. *Developmental Science*, 10(6), 826–835.

Moll, H., Carpenter, M., and Tomasello, M. (2010). Social engagement leads 2-year-olds to overestimate others. *Infancy*, no. doi: 10.1111/j.1532-7078.210.00044.x.

Moll, H., Koring, C., Carpenter, M., and Tomasello, M. (2006). Infants determine others' focus of attention by pragmatics and exclusion. *Journal of Cognition and Development*, 7(3), 411–430.

Moll, H., and Meltzoff, A. N. (2011). How does it look? Level 2 perspective-taking at 36 months of age. *Child Development*, 82(2).

Moll, H., Richter, N., Carpenter, M., and Tomasello, M. (2008). Fourteen-month-olds know what “we” have shared in a special way. *Infancy*, 13(1), 90–101.

Moll, H., and Tomasello, M. (2004). 12- and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science*, 7(1), F1–F9.

Moll, H., and Tomasello, M. (2006). Level 1 perspective-taking at 24 months of age. *British Journal of Developmental Psychology*, 24, 603–613.

# Perspective-Taking and its Foundation in Joint Attention

---

Moll, H., and Tomasello, M. (2007a). Cooperation and human cognition: The Vygotskian intelligence hypothesis. *Philosophical Transactions of the Royal Society B*, 362(1480), 639-648.

Moll, H., and Tomasello, M. (2007b). How 14- and 18-month-olds know what others have experienced. *Developmental Psychology*, 43(2), 309-317.

Moore, C., and Corkum, V. (1994). Social understanding at the end of the first year of life. *Developmental Review*, 14, 349-372.

O'Neill, D. K. (1996). Two-year-old children's sensitivity to a parent's knowledge state when making requests. *Child Development*, 67, 659-677.

Pack, A. A., and Herman, L. M. (2004). Bottlenosed dolphins (*Tursiops truncatus*) comprehend the referent of both static and dynamic human gazing and pointing in an object-choice task. *Journal of Comparative Psychology*, 118, 160-171.

Perner, J. (2000). RUM, PUM, and the perspectival relativity of sortals. In J. W. Astington (ed.), *Minds in the making: Essays in honour of David R Olson* (pp. 212-232). Malden, MA: Blackwell Publishers.

Perner, J., Brandl, J. L., and Garnham, A. (2003). What is a perspective problem? Developmental issues in belief ascription and dual identity. *Facta Philosophica*, 5, 355-378.

Perner, J., Stummer, S., Sprung, M., and Doherty, M. (2002). Theory of mind finds its Piagetian *perspective*: Why alternative naming comes with understanding belief. *Cognitive Development*, 17, 1451-1472.

Piaget, J., and Inhelder, B. (1956). *The child's conception of space*. London: Routledge.

Pillow, B. H., and Flavell, J. H. (1986). Young children's knowledge about visual perception: Projective size and shape. *Child Development*, 57(1), 125-135.

# Perspective-Taking and its Foundation in Joint Attention

---

Racine, T. P., Leavens, D. A., Susswein, N., and Wereha, T. J. (2008). Conceptual and methodological issues in the investigation of primate intersubjectivity. In F. Morganti, A. Carassa, and G. Riva (eds.), *Enacting intersubjectivity: A cognitive and social perspective to the study of interactions* (pp. 65–79). Amsterdam: IOS Press.

Rakoczy, H., Warneken, F., and Tomasello, M. (2007). “This way!”, “No! That way!”—3-year-olds know that two people can have mutually incompatible desires. *Cognitive Development*, 22, 47–68.

Repacholi, B. M., and Meltzoff, A. N. (2007). Emotional eavesdropping: Infants selectively respond to indirect emotional signals. *Child Development*, 78, 503–521.

Repacholi, B. M., Meltzoff, A. N., and Olsen, B. (2008). Infants' understanding of the link between visual perception and emotion: “If she can't see me doing it, she won't get angry.” *Developmental Psychology*, 44, 561–574.

Scaife, M., and Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, 253, 265–266.

Schloegl, C., Kotrschal, K., and Bugnyar, T. (2007). Gaze following in common ravens, *Corvus corax*: Ontogeny and habituation. *Animal Behaviour*, 74, 769–778.

Searle, J. R. (1992). *The rediscovery of the mind*. Cambridge, MA: MIT Press.

Seemann, A. (2007). Joint attention, collective knowledge, and the “we” perspective. *Social Epistemology*, 21(3), 217–230.

Selman, R. L. (1980). *The growth of interpersonal understanding: Developmental and clinical analyses*. New York: Academic Press.

Sodian, B., Thoermer, C., and Metz, U. (2007). Now I see it but you don't: 14-month-olds can represent another person's visual perspective. *Developmental Science*, 10(2), 199–204.

# Perspective-Taking and its Foundation in Joint Attention

---

Southgate, V., Senju, A., and Csibra, G. (2007). Action anticipation through attribution of false belief by two-year-olds. *Psychological Science*, 18, 587–592.

Stekeler-Weithofer, P. (2005). *Philosophie des Selbstbewußtseins: Hegels System als Formanalyse von Wissen und Autonomie*. Frankfurt: Suhrkamp.

Tomasello, M. (2006). Why don't apes point? In N. Enfield and S. C. Levinson (eds.), *Roots of human sociality: Culture, cognition, and interaction* (pp. 506–524). Oxford: Berg.

Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28, 675–735.

Tomasello, M., Carpenter, M., and Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, 78(3), 705–722.

Tomasello, M., and Haberl, K. (2003). Understanding attention: 12- and 18-month-olds know what is new for other persons. *Developmental Psychology*, 39, 906–912.

Tomasello, M., Hare, B., and Agnetta, B. (1999). Chimpanzees, *Pan troglodytes*, follow gaze direction geometrically. *Animal Behaviour*, 769–777.

Wanska, S. (1984). The relationship of spatial concept development to the acquisition of locative understanding. *The Journal of Genetic Psychology*, 145, 11–21.

Wellman, H. M., Cross, D., and Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72, 655–684.

Wimmer, H., Hogrefe, G. J., Perner, J. (1988). Children's understanding of informational access as a source of knowledge. *Child Development*, 59, 386–396.

# Perspective-Taking and its Foundation in Joint Attention

---

## Notes:

(1) We limit our analysis to the first 4 to 5 years of life, thereby not taking account of any higher-order understanding of perspectives that may follow, for example, in adolescence.

(2) Though an implicit ability for level 1 visual perspective-taking may be in place by only 14 months of age, as evidenced by looking-time measures (see Luo and Baillargeon, 2007; Sodian, Thoermer, and Metz, 2007).

(3) It may seem surprising that children solved the Color Mix Task equally well as the Color Task—even though the former involved subtractive color mixing, which 3-year-olds may not know about. But note that prior to the test, children were exposed to the color filters and experienced the color change of the objects themselves (a white object was held behind a blue filter and a blue object behind a yellow filter).



Access brought to you by: MPI fuer Evolutionare  
Anthropologie