(around 8 months of age) in terms of underlying capacities for representation and deduction. The contemporary research on physical and social knowledge in infants (e.g., solidity, gravity, numerical object relations, and implications of desire) that relies on manipulation of expectancy (e.g., use of the habituation/recovery method) is at least implicitly assuming logical processing on the part of the infant. Csibra et al. (1999) argue for a view of young infants employing "principle based" reasoning about rational action when engaged in the teleological or intentional stance.

Implications. Contrary to the simulationist perspective favored by the authors, infants (as logical determinists) may conceivably use evidence of equi-finality and equi-origin to provide a causal understanding of others as intentional. Indeed, it is conceivable that they may do this in coming to understand there own intentionality. If infants can infer mental state variation by virtue of a primitive determinist stance and some primitive capacity for causal logic, then these factors should be examined for phylogenetic and ontogenetic differences. If comparative analysis of a commitment to determinism is as challenging as analysis of inference (Watson et al. 2001), it will not be an easy task.

Triadic bodily mimesis is the difference

Jordan Zlatev,[a] Tomas Persson,[b] and Peter Gärdenfors[b]

Departments of [a]*Linguistics and* [b]*Cognitive Science, Lund University, Lund, Sweden.* **Jordan.Zlatev@ling.lu.se Tomas.Persson@lucs.lu.se Peter.Gardenfors@lucs.lu.se**
url: http://www.ling.lu.se/persons/JordanZlatev.html
url: http://213.80.36.53/temp/lucs/staff/person.asp?id=87&lang=eng
url: http://www.lucs.lu.se/People/Peter.Gardenfors/index.html

Abstract: We find that the nature and origin of the proposed "dialogical cognitive representations" in the target article is not sufficiently clear. Our proposal is that (triadic) *bodily mimesis* and in particular *mimetic schemas* – prelinguistic representational, intersubjective structures, emerging through imitation but subsequently interiorized – can provide the necessary link between private sensory-motor experience and public language. In particular, we argue that shared intentionality requires triadic mimesis.

Tomasello et al. claim that the crucial difference between human beings and apes is "an adaptation for participating in collaborative activities involving shared intentionality" (sect. 6, para. 2), as a result of which human children develop "dialogic cognitive representations." We are sympathetic to this proposal but do not find the nature and the origin of these representations sufficiently clear. Tomasello et al. do not state whether they have an expression-content structure and, if so, whether they are based on language itself, making the account somewhat circular. As for their emergence, it is much too vague to suggest that they are "in some way internalized in Vygotskyan fashion" (sect. 5.2, para. 11).

As an alternative, we suggest that a crucial difference between us and other animals is our advanced capacity for *bodily mimesis*, in particular in its relation to the formation of the *sign function* (Sonesson 1989). Departing from the work of Piaget (1948) and Donald (1991, 2001), we have developed an account of how bodily mimesis provides a "missing link" in both human evolution and child development (Zlatev in press; Zlatev et al., 2005) which we here suggest as an alternative to Tomasello et al.'s account.

Bodily mimesis is based on the ability to match the largely proprioceptively defined *body schema* and more multimodal and consciously accessible *body image* (Gallagher 1995) with perceptions from other modalities. We propose a four-stage model of the development of bodily mimesis that we call the *mimesis hierarchy* and argue that it can be applied to ontogeny as well as phylogeny.

The simplest form, *protomimesis*, involves "body matching," but lacks either clear differentiation between the body schema of the subject and the phenomenon that it is being matched to – paradigmatically the body of the other – or volitional (conscious) control or both. We believe that this is what underlies the neonatal mirroring of human infants (Meltzoff & Moore 1977) and chimpanzees (Myowa-Yamakoshi et al. 2004). The next step is *dyadic mimesis*, which features both differentiation between one's body and whatever it corresponds to and conscious control of the body's representational movements. We consider the capacity for *deferred imitation* to be clear evidence for such differentiation and volition, as well as the capacity for *mirror-self recognition*. (In the latter case, it is the mirror image that is the signifier, while one's own body is the signified.) What these very different skills share is a basic form of the sign function: understanding that one entity can *correspond to* another. Deferred imitation is claimed to be witnessed in human infants as early as 6 weeks (Meltzoff & Moore 1994), but this is rather controversial. At any rate, by 14 months children can perform bodily mimes of actions and events that they have observed previously (Piaget 1948), and around the same time they pass the mirror self-recognition test (Amsterdam 1972). Apes have some difficulties with both of these skills, but it appears that they are indeed capable of dyadic mimesis, reflected in both mirror-self recognition (Lin et al. 1992) and the ability to copy bodily shapes on command (Custance et al. 1995).

The crucial difference is in the next step in the hierarchy, *triadic mimesis*, when the sign function is recruited for communicative means so that an intentional bodily movement, a gesture, stands for an action, event, or object for an addressee. This *communicative sign function* is in essence what is involved in both indexical gestures (i.e., declarative pointing) and iconic gestures. Apes that have not been "enculturated" (cross-fostered) are poor at these skills, though some rudiments have been observed in captive gorillas (Tanner & Byrne 1996) and bonobos (Savage-Rumbaugh et al. 1977). Considering that acts of mimesis can be *interiorized* – that is, performed covertly in imagination – such covert mimetic schemas can be hypothesised to ground the meaning of the child's first words, and in particular the first verbs, the acquisition of which is claimed to be "the major turning point in children's transition to adult-like competence" (Tomasello 1992, p. 7). *Mimetic schemas* are representational and dynamic structures that are accessible to consciousness and *prereflectively shared*. Since they derive from imitating culturally salient actions and objects, both their representational and experiential content can be intersubjective – though not in the strong sense of being mutually known to be shared in the manner of symbols or conventions. What differentiates mimesis from language is above all that the latter is (in principle) fully conventionalized, that is, an object of common knowledge (Clark 1996) and systematic. Thus, signed languages such as American Sign Language constitute the fourth step of the mimesis hierarchy: *postmimesis*.

As sketched above, humans and apes begin to diverge as early as dyadic mimesis, but the difference becomes blatant with triadic mimesis, when mimetic acts and schemas are used for intentional communication. Considering this difference, it is unsurprising to find major differences in apes' and children's capacity for language acquisition. Bodily mimesis is also implicated in intersubjectivity. Gärdenfors (2003, ch. 4) proposes that the capacity for theory of mind be split up in several levels, which fits well with the general approach of Tomasello et al. In particular, we believe that the "shared intentionality" that is the focus of the target article cannot arise without the presence of triadic mimesis. The reason is that the participants can only identify (or create) the shared goal if they can communicate referentially about it, which can be achieved through triadic mimesis.

In conclusion, we hypothesize that the source of human cognitive specificity lies above all in two, mutually enforcing factors: (1) the capacity to form mimetic schemas: representations derived from imitation of public events, and thus basically intersubjective, and (2) the use of mimetic schemas for intentional communication: triadic mimesis. The first of these factors is predominantly representational and the second is communicative, and it is likely that the two have co-evolved. Like Tomasello et al., we consider the crucial human adaptation to be one of social cognition not specifically related to language, but to the motivation and ability

https://doi.org/10.1017/S0140525X05530127 Published online by Cambridge University Press

to form *shared representations*. The difference is that our model, and in particular our construct of mimetic schemas, is more specific. We submit that our model accounts for much of the evidence involving the capacities and limitations of apes, summarized by Tomasello et al., and for the ease with which (nonautistic) children co-construct a world of meaning that is shared with their elders and peers. Thus, mimetic schemas serve as a prerequisite to the acquisition of language, which is just the icing on the cognitive cake, albeit a thick and rich icing.

Authors' Response

In Search of the Uniquely Human

Michael Tomasello, Malinda Carpenter, Josep Call, Tanya Behne, and Henrike Moll

Max Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany. **tomas@eva.mpg.de carpenter@eva.mpg.de call@eva.mpg.de behne@eva.mpg.de moll@eva.mpg.de**

Abstract: As **Bruner** so eloquently points out, and **Gauvain** echoes, human beings are unique in their "locality." Individual groups of humans develop their own unique ways of symbolizing and doing things – and these can be very different from the ways of other groups, even those living quite nearby. Our attempt in the target article was to propose a theory of the social-cognitive and social-motivational bases of humans' ability and propensity to live in this local, that is, this cultural, way – which no other species does – focusing on such things as the ability to collaborate and to create shared material and symbolic artifacts.

R1. Introduction

The 31 commentaries on our target article were for the most part very helpful and thoughtful. We must therefore begin our response, as many authors of responses before us, by apologizing in advance to the commentators for not being able – due to space limitations – to respond with equal helpfulness and thoughtfulness.

R2. Understanding goal-directed action: Who does and who doesn't?

Behaviorists created a scientific paradigm on the fact that we do not observe the goals or psychological states of other organisms directly. Virtually everyone has gone beyond this now – with the current exceptions of **Dymond & McHugh**, who want to construe triadic interactions as "derived relational responding" and **Provine**, who wants to say that intentionality is overrated.

But what about chimpanzees and human babies? Do they understand others in terms of goals and other psychological states? **Povinelli & Barth** think that chimpanzees do not because experiments suggesting that they do (in which they, for example, react differently to purposeful and accidental actions) have other interpretations: the chimpanzees might just be reacting to behavioral cues associated with goal-directed action. This derived behaviorism (the organism has psychological states but doesn't know that others do) is based on a methodological argument that the internal psychological states of an actor can be inferred by an observer only if there is some external expression, and so why not just say the observer is reacting to the external expression and be done with it? One problem is that this argument also applies equally well to human beings, and so if one buys the methodological argument, then it must be assumed that all species, including humans, are equally "mindblind" – which these authors, in other writings, do not claim to be true. Moreover, empirically there are now several different paradigms suggesting that chimpanzees understand that others have goals and see things, and to explain them all without positing a knowledge of psychological states one needs a whole battery of unrelated behavioral cue-reading heuristics based on either innate mechanisms or learning histories that are, to us at least, quite implausible (there is certainly no independent evidence for them). **Watson** would also like to get rid of the understanding of psychological states as a basic cognitive skill, claiming that "intentions and other dispositional states may be inferred logically from an evolved commitment to determinism." In other words, organisms understand only causality, not intentionality, and so Mom's behavior is explained in the same way as a rock's – except that the timing of the contingencies is a bit different. But this cannot be right, as infants understand that people are not happy when their goals are unsatisfied, which must be based on some kind of teleological reasoning in which the actor compares the real state of affairs (as he perceives it) to some desired state of affairs represented internally – a process rocks simply do not engage in.

Perner & Doherty also believe that infants have some kind of externalist way of explaining the behavior of others that does not rely on an understanding of others as goal-directed agents with internal goals. Only when children can understand beliefs, which may be false, can they be said to truly understand the internal mental states of others. The basic argument here is that nothing short of beliefs are considered truly representational because, for example, "goals cannot be misrepresented." But, as first argued by Tomasello (1995), goals can be unsatisfied and the fact that infants understand this – and, again, understand that the actor is not happy when his goals are unsatisfied – is a major piece of evidence that infants understand something about others' internal goals. And goals do have representational content – the actor mentally represents the state of affairs desired – it is just that there is not, as there is in the case of beliefs, a commitment to the "truth" of this representation. But why should truth be the be all and end all? In other words, we simply do not buy the overly restrictive representationalist theory, which says that the only truly mental states are those accompanied by a commitment to their truth, and that knowledge of other representational states such as future-oriented goals is by definition externalistic and non-(meta)representational. **Fernyhough** seems to share some of this minimalist orientation, as he thinks we have overattributed to infants, and that true perspective taking can take place only through the mediation of public representations such as linguistic symbols. We of course agree that some special cognitive abilities emerge later in development with truth-bearing mental representations and linguistic representations (Tomasello & Rakoczy 2003), but we also think that earlier social-cognitive understandings not based on these things are more fundamental and get the process started.

https://doi.org/10.1017/S0140525X05530127 Published online by Cambridge University Press