

## Book Symposium

## Open Access

Michael Schmitz

# A History of Emerging Modes?

DOI 10.1515/jso-2015-0054

**Abstract:** In this paper I first introduce Tomasello's notion of thought and his account of its emergence and development through differentiation, arguing that it calls into question the theory bias of the philosophical tradition on thought as well as its frequent atomism. I then raise some worries that he may be overextending the concept of thought, arguing that we should recognize an area of intentionality intermediate between action and perception on the one hand and thought on the other. After that I argue that the co-operative nature of humans is reflected in the very structure of their intentionality and thought: in co-operative modes such as the mode of joint attention and action and the we-mode, they experience and represent others as co-subjects of joint relations to situations in the world rather than as mere objects. In conclusion, I briefly comment on what Tomasello refers to as one of two big open questions in the theory of collective intentionality, namely that of the irreducibility of jointness.

**Keywords:** Mode; Cooperation; Joint attention; Thought; Propositional attitudes.

## 1 Introduction

Michael Tomasello is a philosopher's dream psychologist. He liberally sprinkles his books with Wittgenstein quotes. He constructively and sympathetically engages with many contemporary philosophers. In his new book *A Natural History of Thinking* he integrates the theoretical fruits of these engagements into an account of the evolution of primate intentionality and thought that is richer and bolder than anything a philosopher would even dare to present. As somebody who has, in addition to his philosophical competence, an overwhelming command of the fields of primatology and of developmental psychology and linguistics, and has made many original contributions to them, he is perhaps uniquely positioned to do so. How could a philosopher try to respond to this

---

**Michael Schmitz**, Department of Philosophy, University of Vienna, Vienna 1010, Austria,  
e-mail: Michael.Schmitz@univie.ac.at

 ©2016, Michael Schmitz, published by De Gruyter.

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 3.0 License.

work in the spirit and mode of cooperation that Tomasello's book so incisively analyzes, so richly documents, and so generously embodies? As the title of the book already indicates, Tomasello makes a point of focusing on the nature of human thought, and this emphasis is also part of what distinguishes the approach of this book from that of some of his earlier work as, e.g. *The Cultural Origins of Human Cognition* (2009). Another major difference is that he now sees the defining characteristic of human thought exclusively in its social, cooperative nature. So in this commentary I will focus on Tomasello's understanding of the nature of thought and on how it relates to traditional ways of understanding thought in philosophy.

In the broadly Cartesian tradition that Tomasello responds to, there is a deep-seated tendency to identify consciousness with thought. Thought in turn has recently usually been thought of as a "propositional attitude." This tradition has a profound intellectualist bias in that it tends to disregard non-propositional, non-conceptual forms of thought and consciousness. It is also theory-biased (Schmitz 2013a,b) in the sense of privileging theoretical, mind-to-world direction of fit attitudes or speech acts – postures – like beliefs and assertions, over practical, world-to-mind direction of fit postures like intentions and orders. This dominance of the theoretical in turn is manifest in at least two ways. First in the idea that all propositional attitudes, whether practical or theoretical, are attitudes towards (or contain) something that, as a truth value bearer, belongs to the theoretical domain – as truth is representational success from a theoretical rather than from a practical point of view. Second, propositional attitudes are commonly approached by analyzing their reports – such as he believed this, she intends that, and so on – and that means as objects of a theoretical perspective that takes them to be something that is the case. The subject of the attitude and its kind or mode are merely reported from a theoretical perspective and are not thought to contribute to its content, which is taken to be identical to that of the relevant proposition. Finally, the received notion of a propositional attitude also shows an individualist bias in that it is usually taken for granted that the relevant subjects are individuals.

*A Natural History of Human Thinking* relates to all of these (and some further) aspects of the philosophical tradition on thought in interesting ways. In the following I will first introduce Tomasello's notion of thought and his account of its emergence and development through differentiation, arguing that it calls into question the theory bias of the tradition as well as its frequent atomism. On the basis of this discussion, I then raise some worries that he may be overextending the concept of thought, arguing that we should recognize an area of intentionality intermediate between action and perception on the one hand and thought on the other. I continue this argument by suggesting that the cooperative nature

of humans is reflected in the very structure of their intentionality and thought: in cooperative modes such as the mode of joint attention and action and the we-mode, they experience and represent others as co-subjects of joint relations to situations in the world rather than as mere objects. To capture this I suggest to abandon the traditional understanding of propositional attitudes. In conclusion, I will then briefly comment, in the light of the preceding discussion, on what Tomasello refers to as one of two big open questions in the theory of collective intentionality, namely that of the irreducibility of jointness.

## 2 The Nature, Emergence and Development of Thought

Inspired by Ronald Langacker's (1987) conception of linguistic schemas, Tomasello views thought essentially as a schematization of and abstraction from experience. He focuses on three key components of thought: representation, self-monitoring, and inference. "Representation" for Tomasello essentially involves an "ability to cognitively represent experiences to oneself 'offline'" (p. 3). The notion of "self-monitoring" is especially meant to highlight the evaluation of behavioral outcomes that thoughts lead to (p. 3). The emphasis on inference indicates that Tomasello's notion of thought is more process-oriented than the state-oriented mainstream philosophical tradition. An even more important difference is that Tomasello allows for the possibility of thought that is situational in that it is directed at whole situations rather than just things, properties, or relations, but yet not propositional because it represents these situations in an iconic, imagistic, rather than a propositional representational format. In doing this, Tomasello takes an important step beyond the intellectualism of the tradition.

Tomasello's approach also provides a convincing alternative to the frequent atomism of the tradition. Cognitive evolution does not proceed from simples to ever more complex wholes – regardless of whether these simples are the stimulus-response linkages of the behaviorist or the simple ideas of the mentalist – "...but rather from inflexible adaptive specializations of varying complexities to flexible, individually self-regulated intentional actions" (p. 26). Articulated thought arises through processes of differentiation from (relatively) undifferentiated wholes. Individuals learn to "extract their components and use them in productive combinations" (p. 28).

An example for this differentiation process is provided by the development of force indicators. Tomasello hypothesizes that "early humans' first acts of cooperative communication were pointing gestures in joint collaborative

activities” (p. 50) underlain by communicative motives not yet differentiated between requestive, practical and informative, theoretical forces. An early human pointing out a predator, a prey, or, as in Tomasello’s example, a stick that could be helpful in a joint activity of gathering honey, is neither only drawing attention to something that is the case, nor solely requesting or directing a future course of action. In the immediate context of this situation and the joint activity that imbues it with meaning, these functions will not necessarily be differentiated yet, neither in the communicative act nor in the underlying mental state.

This idea is a useful antidote against what above I called the “theory bias” in many accounts of thought and intentionality. One rarely stated motivation for this bias might be expressed as follows: “Where could the material for thought come from, if not from what is the case? So theoretical thought about what is the case must be more fundamental than thought guiding and directing action.” But as the examples show, there are ways of drawing attention to elements of reality which at the same time direct action, and which are phylogenetically and ontogenetically prior to clearly differentiated theoretical and practical positions towards situations.

The emphasis on differentiation should of course not be taken to mean that integration of elements is not also an essential part of cognitive evolution. Properly understood, differentiation and integration are often complementary, as Tomasello shows when describing the integration of pointing and pantomiming gestures into multiunit expressions. When a gesture pantomiming snake motion and a gesture of pointing to a cave are combined into a single communicative act warning of snakes in the cave, this is integration. But relative to an earlier developmental stage, where broadly the same communicative effect was achieved just by pantomiming, it is also an articulation, a parsing of that act into differentiated elements (p. 67). Corresponding remarks apply to stages of the development of spoken language, such as when pointing gestures are integrated with spoken language, and one-word utterances are differentiated into articulated clauses.

From these and many other observations and theoretical remarks in Tomasello’s book I think we can extract a number of parameters to describe the direction of the evolution of intentionality: from perception, action and interaction in the immediate perceptual context over pointing and pantomiming gestures to spoken and finally written language – and to the corresponding forms of thought. This evolution leads to forms of intentionality that: are ever more independent from an immediate shared perceptual context, ever more removed from the flow of joint sensory-motor activity; are ever more general, abstract and what Tomasello refers to as ‘perspectival’, that is as involving a specific conceptual construal of an entity (e.g. as an animal as opposed to a human, a man, etc.); go from perceptual and actional over deictic and iconic to symbolic forms of representation;

are ever more differentiated and articulated, in particular show a greater degree of differentiation of representational role as e.g. in the increase of different grammatical categories (see Schmitz 2012). Moreover, as Tomasello shows, these stages are connected to characteristic social units with typical sizes ranging from the dyads of second-personal joint attention and action over small band and tribes to nation states. And they are also connected to corresponding forms of normativity, by which these groups create, maintain and extend their joint practices, skills, customs, traditions, sensibilities, norms, rules, laws, institutions, and so on. These in turn range from emotional responses in the immediate context over informal to legal sanctions. A final parameter I want to mention here is the differentiation of roles within a society. The representational richness of the English language could only have been produced by a society (or societies) with a multitude of institutionalized roles from lorry driver to lexicographer. Humans are very role-oriented, and as Tomasello points out, they may even be uniquely equipped to understand role concepts (p. 42).

### 3 The Boundaries of Thought

Whereas Descartes notoriously denied any thought to animals, Tomasello ascribes thought rather generously – not only to primates, but even to squirrels (p. 14) – based on his more liberal conception that allows for pre-propositional, imagistic forms of thought. And I think he is right against the language-centric tradition in philosophy and psychology that, e.g. simulating available courses of action and weighing them against one another is naturally described as thinking. That thought requires a propositional representational format is just a dogma. However, at certain points I also felt that Tomasello may be overgenerous in ascribing thought processes.

Let me here focus on one of Tomasello's examples for thought processes in great apes, based on Buttelmann et al.'s (2007) test of six human-raised chimpanzees in the so-called "rational imitation paradigm" (Gergely et al. 2002):

Individuals saw a human perform an unusual action on an apparatus to produce an interesting result. The idea was that in one condition the physical constraints of the situation forced the human to use an unusual action; for example, he had to turn on a light with his head because his hands were occupied holding a blanket, or he had to activate a music box with his foot because his hands were occupied with a stack of books. When given their turn with the apparatus and no constraints in effect, the chimpanzees discounted the unusual action and used their hands as they normally would. However, when they saw the human use the unusual action when there was no physical constraint dictating this – he just turned

on the light with his head for no discernable reason – they quite often copied the unusual behavior themselves (p. 23).

Tomasello goes on to suggest that the “most natural interpretation” of the response pattern is that the apes employed an imagistic, pre-propositional “*proto-modus tollens*”: “(1) he is not using his hands; (2) if he had a free choice, he would be using his hands; (3) therefore he must not have a free choice” (p. 23).

But does it really take inference and thought to understand what is going on and to explain why the apes react in the way they do? I want to suggest we can make sense of their response pattern while remaining at a sensory-motor level. The chimpanzees perceive and understand what the human is doing. The kind of understanding I have in mind here is enactive. It may just be manifest in imitation, in completing the relevant action, or – if we are dealing with more cooperatively inclined beings such as humans – in helping others to perform it. It does not imply that they ascribe intentions or other mental states to the human. But it does mean that they are able to understand what is essential to the action and what is an adaption to unusual circumstances. Their imitation behavior manifests this understanding. They perceive the fact that the human is carrying a blanket as a constraint on his actional capacities and show that they understand that this is why he uses his head by not copying this aspect of his action in a context where they are not subject to the same constraint. It is inessential to the action of turning on the light.

In contrast, when the unusual manner of performance cannot be attributed to unusual circumstances, it is experienced as surprising, as deviating from the schema for this action. I would suppose that the apes experience this behavior as intriguing. It piques their curiosity, and so in this case they imitate the manner of action because it seems to be essential to the action performed and because they want to understand what its point is. It is tempting to gloss what is going through the chimpanzees’ minds e.g. as follows: “I wonder why this human is behaving in this weird way. What is the point, what is he doing? Perhaps it is especially rewarding, fun, or otherwise interesting, so I want to try it, too!” and then to continue as follows: “This is – of course – not exactly what is going through the ape’s (or the small child’s or even the adult’s) mind. But it is the explicit form of what is being thought implicitly.” This kind of appeal, popular as it is, strikes me as no more than convenient handwaving: we are in effect saying that what is actually going through the subject’s mind is in some ways like the propositional form of thought and in some ways not, but we are not really saying anything substantive beyond that. I want to suggest in a Wittgensteinian vein that, even though we might get more detailed, we may have pretty much already given the outline of what is going on: in the first set of

cases they experience and understand the unusual behavior as a response to the unusual circumstances and therefore do not copy it, but in the second set of cases they cannot: they are puzzled and intrigued by it and therefore go on to imitate it. In similar cases this is all that has gone through my mind and I believe that of other humans. This may just be bedrock – at least on the level of intentionalist explanation.

Of course, additional things might be going on. For example, the chimpanzees may very well be actively scanning the human and the broader situation, trying to figure out what is going on. But searching for clues in this way is not yet thinking; it is not sufficiently decontextualized and entirely manifest in action and perception. Of course they might also be simulating – as Tomasello explicitly suggests for other examples. For example, they might be simulating what it is like to turn on the light with their heads to understand the point of this action before actually performing it. But I do not see why this would necessarily be the case. Consistent with the spirit of Tomasello's general outlook and the broad pattern of development described above, we would expect a sequence where puzzlement over the action would first immediately lead to overt imitation, and only later overt imitation would be preceded by its internalized counterpart: simulation in imagistic thought.

This example and others, some of which I will soon discuss, raise the general worry that Tomasello may be somewhat overextending the concept of thought. Or, to put it differently, he may be underestimating what can be explained by refined actional and perceptual skills and the *sense* subjects, including animal subjects, may have of what is important and what to do in a given situation. For the most part, he appears to be operating with a restricted set of alternatives: that an action should be explained in strict behaviorist or associationist terms, or else by thought.<sup>1</sup> But if we allow ourselves to ascribe to some animals the capacity to abstract and schematize things in thought, I think we can also explore the idea that they pick out, extract, or schematize certain features in experience – for example, what is essential to an action and what not – and manifest this understanding in action. The difference would be that these capacities might be entirely realized in the attentional structure of sensory-motor-flow, in refined actional and perceptual sensibilities, in searching and orienting behavior; but also in certain broadly emotional aspects of experience such as feelings of surprise, familiarity, curiosity, amusement, joy, fear; in senses, hunches or instincts that an action is the right one (or not), confidence or its absence in performing it, and so on. With

---

<sup>1</sup> There is only a passing reference to 'intuitive heuristics' and Kahnemann at the beginning (p. 3) which is not taken up again later.

a nod to Piaget, though without taking on board all elements of his theory, we might speak of “sensory-motor-emotional schemata” here (Schmitz 2012).

It seems to me that much is going on here below the levels of stepping out of the sensory-motor flow, of reflecting, of internalizing action and perception in imagination and then in propositional forms of thought. The point is not where exactly the boundary between thought and sensory-motor-emotional experience should be drawn, nor even that a sharp boundary should be drawn at all. The boundary is not likely to be sharp, and rather than trying to sharpen it, it may be more useful to think about this just in terms of a continuum of the parameters discussed above (and possibly others). The point is that even if we go along – as we should – with Tomasello’s more liberal conception of thought, there is still a large territory to be explored between thought and what can be captured in strict behaviorist, associationist or similarly restrictive terms, and that explanations for some of the phenomena discussed by Tomasello may come from there rather than from thought.

## 4 The Cooperative Nature of Human Thought and Intentionality

I now want to extend this point to human cooperative intentionality. A striking illustration of the cooperative mindset of humans in comparison to even our nearest living relatives, the great apes, comes from the so-called “object choice paradigm”: whereas human children as young as 12 months old immediately understand that a human experimenter who points to one of two containers that may have food in them wants to help them by indicating the one with the food, apes cannot interpret this gesture and indeed seem rather dumbfounded by it – it does not even occur to them that the human might want to help them. Why do humans respond so differently? According to Tomasello the “mutual assumption of cooperativeness in such situations is so natural for humans that they have developed a special set of signals – ostensive signals such as eye contact and addressing the other vocally – by means of which the communicator highlights for the recipient that he has some relevant information for her” (p. 52). Tomasello is certainly right that this is because humans have a much more cooperative mind-set than apes. But what does it mean that they make a “mutual assumption of cooperativeness”? Surely Tomasello does not mean that the infants actually think something like “This is a human. Humans are generally cooperative, so this human is probably trying to help me by pointing to the bucket. So that could be where the food is.” And again, I believe we should not

be content with just saying something to the effect that the infants are thinking this implicitly.

Perhaps Tomasello has in mind something similar to John Searle's (e.g. 1992, 1995) notion of a background assumption, roughly something we take for granted without thinking about it, at least not normally. Searle thinks of the background as being non-intentional in the sense of being non-representational, but this means that this talk about assumptions must ultimately be metaphorical: how could something be an assumption without being a possible content of thought and how could something with a specific content fail to be intentional (Schmitz 2012)? So I think it is hard to make sense of the notion of a background assumption in general and in particular of Tomasello's notion of an assumption of cooperativeness *at the level of understanding that we are concerned with here* – of course there are levels where people do make such assumptions. To count as making an assumption we need at least to be able to think the content of that assumption in thought, and it is implausible that young infants are capable of this. I therefore want to propose to replace that notion at this level through that of a mode of cooperativeness – the mode of joint attention and action – and of having a sense of others as potential cooperation partners.

The suggestion is that the cooperative mindset of infants is primarily manifest in the perceptual, actional and emotional dispositions and sensibilities they have with regard to other creatures – people and animals. They tend to experience others and have a sense of them as potential cooperation partners, and when they do cooperate with them they function in a cooperative mode. Again, this is of course not meant to suggest that the concept of a cooperation partner is applied in their experience. It is rather that they are attuned to people in such a way that in general they tend to trust them to help them and are ready to help them in turn – of course there will always be exceptions in individual cases, unfamiliar contexts, various individual and social pathologies, and so on. To put this in Gibsonian jargon, they will experience them in emotionally charged ways that afford cooperative interactions of various kinds. Even a 1-year-old infant already has experienced countless sensory-motor-emotional interactions of being fed, taken care of, of playing games, exchanging eye-contact, etc. These experiences dispose the infant to expect others to be cooperative and to understand their behavior and respond to it accordingly.

It is tempting to suppose that if one expects something one must believe that it will take place. But here by “expectation” I only mean that one is habituated to certain patterns, that one experiences them as familiar, and is therefore disposed to respond to them in certain ways and to be surprised when there are deviations from them. It is not necessary to have beliefs/thoughts about something and concepts of it to be surprised by it. For example, people are habituated to others

keeping a certain distance to them in normal conversation. Now there are differences between cultures with regard to that distance. This may lead to a member of one culture – say an Italian – moving closer so that the member of another culture – say an Englishman – feels a bit uncomfortable and thus negatively surprised and starts to move away, which causes the Italian to feel uncomfortable so that he moves closer again, and so on.<sup>2</sup> Such an episode may cause one or both of them to think about this and perhaps even to introduce a concept for it like the concept of personal space. But this will not necessarily happen. The two may just experience a slight and undifferentiated discomfort and never think about it and its causes, perhaps because they are too focused on the content of their conversation. So I think we cannot infer conceptual competence just from some kind of dishabituation as is often assumed in the context of the violation of expectation paradigm. Surprise at something is not an index of the presence of corresponding concepts, thoughts and beliefs, but just a possible cause of them. There are countless things that we experience as familiar and that we are habituated and attuned to, but that we only start to think about when there are deviations from these patterns and when we already have sufficient thinking skills – which infants often do not have – and sometimes not even then.

The general suggestion then is that the human infants in the object choice paradigm and many other similar contexts interpret the pointing gesture correctly because they are habituated and attuned to cooperative behavior – which is reflected in their whole sensibility, their actional, perceptual, and emotional experience – and not because they literally make an assumption about cooperation.

## 5 A Mode Account of Human Cooperative Intentionality

Given the convincing case Tomasello makes that humans are profoundly different from other animals in this regard, we may expect this difference to be reflected in the very structure of their intentionality. In this section I want to make a friendly suggestion on how this might be the case. To do this, I need to go back to the received notion of propositional attitudes briefly described above. Recall that on that notion, the representational content of a state is identical to that of the proposition that is its object/content. That is, the subject of, for example, a perceptual or actional state, a belief or an intention, only represents what (s)he perceives,

---

<sup>2</sup> I owe this example to a lecture by John Searle, who observed such a situation at a conference.

does, beliefs or intends. In the spirit of philosophers such as Immanuel Kant and psychologists like Jean Piaget I think this picture should be replaced with one where subject consciousness and object consciousness are seen as interdependent and as the two poles of any intentional state. (This is also broadly in the spirit of Tomasello's emphasis on self-representation as one of the three key components of thought.) I think we are never only aware of a state of affairs, but also always – though typically in a backgrounded sort of way – of our position with regard to this state of affairs and thus of ourselves. For example, in perception I experience the world as acting causally on me and in action I experience myself as acting causally on the world; in intending something I must have at least a sense of that practical position, a sense that, for example, it is up to me to bring about the intended action. This does not mean that I need to have a concept of the relevant position, such that I would be able to think about it. My sense of my position may simply be manifest in the strength of my desire or of my confidence to reach the goal, or in the epistemic feeling of confidence (or lack thereof) that something is the case.

This revised understanding of intentional attitudes now also enables us to account for the special way in which our cooperation partners are represented and figure in our experience. The proposal is that they are represented as co-subjects of joint positions vis-à-vis the world rather than as objects of such positions, as what is believed, intended, perceived or done. For example, in an episode of joint attention, my co-attender is not what I attend to, but who I attend with. Jointly intending something is not a matter of me intending it and believing that you intend it and believe that I intend it and believe that you believe that I intend it, and so, but of both of us having we-attitudes of the form “We intend to...”. Note that it is only the revised understanding of intentional states that makes it possible to accept this simple picture at face value, because on the traditional understanding the subject is singular and only ascribed from an external perspective, as part of a report, so that it is only represented by the reporter, but not by the subject of the attitude itself.

The suggestion then is that the special cooperative nature of the human mind is manifest in the structure of intentional states in that humans can and often do represent others as co-subjects of joint positions or relations to state of affairs rather than as mere objects of individual positions, as part of these states of affairs, whereas non-human animals are much less capable of this or – as according to Tomasello – not even at all. But what does it mean to represent somebody as a co-subject? To give a short answer to this question, let me borrow a metaphor and a basic concept from Ronald Langacker (1987). The metaphor is that to construe somebody subjectively is to construe her as part of one's perceptual apparatus, as part of what gives one access to the world as opposed to what

one accesses, somewhat in the way one experiences one's glasses. Normally one does not attend to one's glasses as objects in their own right but as something that improves one's access to the objects one attends to. We can add that analogously a tennis player is not normally attending to her racket, but experiences it as an extension of her actional apparatus, as something that improves her actional reach in the world. I hasten to clarify that these metaphors fall short in one important respect: while these tools just *serve* my theoretical and practical needs and goals and are in that sense mere objects, my co-subjects share at least some of them. I experience them as being as like me (Meltzoff 2007), as exhibiting what Tomasello often refers to as "self-other equivalence" (e.g. p. 29, 41). The theoretical and practical capacities and positions we share are part of the ground in Langacker's technical sense – roughly, the elements of the speech situation relative to which everything talked about is situated – which thus becomes a common ground.

In experiencing somebody as a co-subject, I do not experience him as something that is the case, for that would be to experience him from a theoretical position, as an object of observation. I experience him as a potential or actual partner for theoretical, epistemic as well as for practical cooperation; as a source of information about the world and at the same time as somebody who will help and guide me; as somebody who draws my (our!) attention to new, exciting, interesting things and who I in turn want to show interesting things to; but also as somebody whom I can trust in a dangerous situation (e.g. social referencing). Note that while these various aspects are conceptually distinguished in the description, they are still undifferentiated, inchoate, at the gestaltlike, not yet conceptually articulated, experiential level. In particular, the practical and theoretical aspects of the co-subject experience are undifferentiated similar to how, as we saw, practical and theoretical aspects are undifferentiated in early pointing gestures.

These points are closely related to the intimate connection between practical and theoretical aspects of jointness emphasized by Tomasello: "Joint actions, joint goals, and joint attention are thus of a piece, and so they must have coevolved..." (p. 44). His argument is that humans must coordinate their attention in order to act jointly. Against certain tendencies in philosophy to think of joint attention as a purely perceptual phenomenon, this argument can also be made in the opposite direction: joint attention can only be joint if the co-subjects are at least disposed to joint action, for nothing else could distinguish it from mere mutual perception (Schmitz 2015). To experience somebody as a co-attender is to experience him as at least a potential co-subject of both theoretical and practical positions towards the world.

This picture of the experience of co-subjectivity in the mode of joint attention and action can be supported and concretized by many results from developmental

psychology, particularly from the study of autistic children, whose capacity for joint attention and action is well-known to be diminished. To mention just a few examples, when asked where a sticker should go, more than half of the children with autism never indicated the place by pointing to their own bodies rather than at the other's body, whereas all neuro-typical children did (Hobson and Meyer 2005). This is a very vivid illustration of the difference between a co-subjective and an objectifying style of reference. To point to a place on one's own body to pick out the corresponding place on that of the other, is to treat her as somebody like oneself rather than as an object. Autistic children also engage much less in the kind of affirmative nodding people often show when listening to others (Hobson et al. 2009). A straightforward interpretation of this is that they experience common ground less and/or have less interest in emphasizing it and in maintaining the emotional connection that it brings with it. Similarly, autistic children also have difficulties "in sustaining engagement with the questioner and appreciating how their communication established common ground between self and other in relation to which a third party was "he" according to a joint perspective" (Hobson and Hobson 2011, p. 118). That being in the joint attention mode is not only manifest in how co-attenders are experienced, but also in how the world is experienced – namely relative to their interests and to the common ground – is illustrated by another finding: non-autistic children were much more likely to show a concerned 'checking' look at a tester, with whom they were in a joint attention situation, when the tester's drawing was torn, than autistic children (Hobson et al. 2009).

So in the mode of joint attention and action, the co-attender is experienced as (1) like me in important respects and as (2) a co-subject of a joint intentional relation to a state of affairs. This relation of sharing a common ground is (3) affectively charged in such a way that (4) the co-subjects are at least disposed to joint action and (5) to experience the world with regard to the theoretical and practical needs, concerns and interests of the co-attender. That experiencing an intentional relation as co-subjects is essentially different from just experiencing it as an object of a single subject can be supported by a study which showed that 14-months-old infants understood an ambiguous request by an adult on the basis of a shared joint attention episode, but not by merely observing his otherwise identical interactions with the relevant objects. After the adult and the infant had shared two objects and the infant had explored one object alone, the infant was able to correctly interpret an ambiguous request for "that one", made with an excited expression by the adult, as referring to the new object. But 14-month-old infants were not able to do the same in conditions where infants merely observed the adult examine the objects by himself, or the adult engaging in joint attention with another person (Moll et al. 2007).

In the book, Tomasello discusses several experiments where the common ground is constituted more through joint action than through joint attention. These can analogously be explained through co-subjectivity. As an example, consider a study by Liebal et al. (2009), who

... had a one-year-old infant and an adult clean up together by picking up toys and putting them in a basket. At one point the adult stopped and pointed to a target toy, which the infant then cleaned up into the basket. However, when the infant and adult were cleaning up in exactly this same way, and a second adult who had not shared this context entered the room and pointed toward the target toy in exactly the same way, infants did not put the toy away into the basket; they mostly just handed it to him, presumably because the second adult had not shared the cleaning up game with them as common ground (p. 55).

So on the present proposal this talk of the common ground can be further explained by saying that the first adult was experienced as a co-subject of the joint action of cleaning up, but the second was not, so that his point was not interpreted in terms of the joint action frame or schema. Co-attenders and cooperators are bound into the representation of the intentional relation to the shared object or goal in a special way – as co-subjects – that is reflected in the very structure of the relevant intentional states.

This general approach is also useful for understanding collective intentionality that is clearly on the level of thought such as joint intention and belief. Since Schelling (1960) and Lewis (1969) it has often been held that coordination requires recursive mindreading of the form: I believe that p and I believe that you believe that I believe that p, and so on, ad infinitum. This is at least partly an artifact of allegiance to the traditional model of propositional attitudes according to which, as we noted above, the subject of the attitude is an individual and neither the subject nor its position are represented. This is what generates the infinite regress because each position an individual takes will be outside of what is represented, so that we need another attitude to represent this, which will again not be represented, and so on, ad infinitum. In this way, we can never capture the meeting of minds, the joint epistemic possession of a state of affairs in joint belief or the joint practical attitude towards an action in joint intention. By contrast, on the revised understanding we can just straightforwardly think of each of the co-subjects as representing their joint position by thinking “we believe that...” or “we intend to...”. As Tomasello puts it: “...human individuals are attuned to the common ground they share with others, and this does not always involve recursive mindreading” (p. 38). The basic idea can further be extended to account for the importance of roles in human cooperation, which Tomasello emphasizes in many places in the book. In “role mode” (Schmitz 2013a) individuals and groups take positions from the point of view of the institutional roles they occupy:

“As chancellor, I believe/intend/have the power...”, “As members of the committee, we believe, intend, demand...”. Different subjects occupy a multitude of roles in the joint pursuit – with their co-subjects in the organization – of certain shared theoretical and practical purposes: another example of the “dual level structure” of a shared objective and differing roles and perspectives with regard to this objective that Tomasello highlights at various points (e.g. p. 43–48).

## 6 Emerging Modes of Cooperation and their Reducibility

As a final point, let me address one of the two big open questions Tomasello takes up towards the end of his book: namely the question of the irreducibility of jointness, collectivity or “we-ness.” Can collectivity be captured in terms “of the individuals involved, and what is going on in their individual heads” (p. 152)? According to many theorists, in particular those who defend ‘relational’ accounts, for example, of joint attention (Campbell 2002), this cannot be done. Tomasello rightly insists that we may still “...ask the evolutionary or developmental question of what does the individual bring to the interaction that enables her to engage in joint attention in a way that other apes and younger children cannot” (p. 152). I think we must go even further and ask what it is about these individuals synchronically – not only diachronically – that joins them together in relations of co-attending, co-intending, co-believing, and so on. (And I think Tomasello in fact does give an answer to this question, namely in terms of the characteristic modes of thought humans employ in cooperation and communication.) There is nothing reductionist about maintaining that these relations obtain in virtue of cooperation-specific intentionality in the minds (and heads) of the individuals involved – whether in the form of the mode account I have suggested or in a different way. Nor is there any reason to deny, implicitly or explicitly, that such intentionality can – like all intentionality – sometimes misrepresent as in illusions of attending jointly or in we-intending goals that the putative co-subjects have already abandoned.

The truly reductionist attitude of course is the one – most ably defended by Michael Bratman (2014) – according to which at least small-scale cooperative activity can be explained with reference to essentially the same kind of intentional capacities as individual action. I believe that this approach is strongly undermined by Tomasello’s argument in *A Natural History*, because he makes a very convincing case that the capacities of great apes for individual action and causal understanding are very similar to those of human infants at a certain stage

of early development, while their capacities for cooperation are much weaker and perhaps even non-existent.

My main critical point has been that Tomasello may underestimate the extent to which these differences in behavior can be explained through differences in sensory-motor-emotional sensibilities rather than through differences in thought. I think we may have to move even further away from the strict traditional dichotomy between thought in the sense of propositional attitudes on the one hand and what can be explained in strict behaviorist or associationist terms than Tomasello already does by embracing imagistic forms of thought. Understanding an action, having a sense of somebody as a cooperation partner, or being in a mode of joint attention and action all do not neatly fit into these categories. And yet these forms of intentionality may help to explain both why the “we” is irreducible and why there is nothing mysterious about collective subjects: because they just are individuals as mutually connected through their intentionality. At the lowest level these connections are made through experiencing others as co-subjects of relations of joint attention and action and through the emotional bond that brings with it. Against the background of these connections we may then begin saying “we” in that special, affectively charged way, in which it has an irreducibly cooperative meaning. In the we-mode people take the connection to their co-subjects to the next level by committing themselves to joint beliefs, plans, values, and so on. Another milestone in the history of human thought may be the ability to relate to other people and to the world from the vantage point of roles defined in institutional contexts relative to the roles of one’s co-subjects in the pursuit of certain shared purposes. The history of human thinking so richly and boldly conceived by Tomasello may also be a history of emerging modes of cooperation.

**Acknowledgments:** I am indebted to Henrike Moll for helpful discussions and comments on a draft of this paper.

## Bibliography

- Bratman, M. E. (2014): *Shared Agency: A Planning Theory of Acting Together*. Oxford: Oxford University Press.
- Buttelmann, D., M. Carpenter, J. Call, and M. Tomasello (2007): “Enculturated Chimpanzees Imitate Rationally”. In: *Developmental Science* 10. No. 4, p. 31–38.
- Campbell, J. (2002): *Reference and Consciousness*. Oxford: Clarendon Press.
- Gergely, G., H. Bekkering, and I. Király (2002): “Developmental Psychology: Rational Imitation in Preverbal Infants”. In: *Nature* 415. No. 6873, p. 755–755.
- Hobson, J., R. Harris, R. García-Pérez, and R. P. Hobson (2009): “Anticipatory Concern: A Study in Autism”. In: *Developmental Science* 12. No. 2, p. 249–263.

- Hobson, R. P. and J. Hobson (2011): “Joint Attention or Joint Engagement? Insights from Autism”. In: A. Seemann (Ed.): *Joint Attention: New Developments in Psychology, Philosophy of Mind, and Social Neuroscience*. Cambridge: MIT Press, p. 115–136.
- Hobson, R. P. and J. A. Meyer (2005): “Foundations for Self and Other: A Study in Autism”. In: *Developmental Science* 8. No. 6, p. 481–491.
- Langacker, R. W. (1987): *Foundations of Cognitive Grammar: Theoretical Prerequisites* (Vol. I). Stanford: Stanford University Press.
- Lewis, D. (1969): *Convention: A Philosophical Study*. Cambridge: Harvard University Press.
- Liebal, K., T. Behne, M. Carpenter, and M. Tomasello (2009): “Infants use Shared Experience to Interpret Pointing Gestures”. In: *Developmental Science* 12. No. 2, p. 264–271.
- Meltzoff, A. N. (2007): “‘Like me’: a Foundation for Social Cognition”. In: *Developmental Science* 10. No. 1, p. 126–134.
- Moll, H., M. Carpenter, and M. Tomasello (2007): “Fourteen-Month-Olds Know What Others Experience Only in Joint Engagement”. In: *Developmental Science* 10. No. 6, p. 826–835.
- Schelling, T. C. (1960): *The Strategy of Conflict*. Cambridge: Harvard University Press.
- Schmitz, M. (2012): “The Background as Intentional, Conscious, and Nonconceptual”. In: Z. Radman (Ed.): *Knowing without Thinking: Mind, Action, Cognition and the Phenomenon of the Background*. Basingstoke: Palgrave, p. 57–82.
- Schmitz, M. (2013a): “Limits of Intention and the Representational Mind”. In: G. Seebass, M. Schmitz, and P. M. Gollwitzer (Eds.): *Acting Intentionally: Individuals, Groups, Institutions*. Berlin, New York: DeGruyter, p. 57–84.
- Schmitz, M. (2013b): “Practical Knowledge”. In: Miguel Hoeltje, Thomas Spitzley und Wolfgang Spohn (Eds.): *“Was sollen wir glauben? Was dürfen wir tun?” Sektionsbeiträge der GAP.8*. DuEPublico, p. 392–403. Retrieved from [http://duepublico.uni-duisburg-essen.de/servlets/DerivateServlet/Derivate-33085/GAP8\\_Proceedings.pdf](http://duepublico.uni-duisburg-essen.de/servlets/DerivateServlet/Derivate-33085/GAP8_Proceedings.pdf).
- Schmitz, M. (2015): “Joint Attention and Understanding Others”. In: *Synthesis Philosophica* 58, p. 235–251.
- Searle, J. R. (1992): *The Rediscovery of the Mind*. Cambridge: MIT Press.
- Searle, J. R. (1995): *The Construction of Social Reality*. New York: Free Press.
- Tomasello, M. (2009): *The Cultural Origins of Human Cognition*. Cambridge: Harvard University Press.