

Book Symposium

Open Access

Michael Tomasello

Response to Commentators

DOI 10.1515/jso-2015-0042

Abstract: This paper is a reply to the comments by Henrike Moll, Glenda Satne, Ladislav Koreň and Michael Schmitz on Michael Tomasello, *A Natural History of Human Thinking* (Harvard University Press, 2014).

Keywords: Thinking; Shared intentionality; Collective intentionality; Normativity; Evolution; Perspective.

I would like to thank the four commentators very much for their time and thoughtful commentary. In writing the book I made a conscious effort to make contact, wherever I could, with philosophical modes of discourse (albeit as an outsider). The responses here show that at least some contact was made.

Nevertheless, there were some misunderstandings, due mostly to my use of some key terms in ways that are not conventional in the philosophical community. For example, Koreň points out a supposed inconsistency: I attribute “knowledge” to chimpanzees when they are not supposed to understand beliefs, the assumption being that knowledge is justified true beliefs. Although I might not have been as explicit as I could have been, and so this is my fault, in many other places I have been clear that all that is meant by “knowledge” is that chimpanzees know when a competitor “has seen” something happen (and will react knowing that the other has “registered” the event). In philosophical jargon, it is knowledge by acquaintance. Something similar is true of the word “intention”, which can be used in two fairly different ways. In previous publications (e.g. Tomasello et al. 2005) we defined intention as a plan toward a goal (a desired state of affairs) to which the individual was committed. But then there is the larger notion of intentionality involving propositional representations, which my use of the word intention can sometimes be confused. Without being fully explicit about it, I have tended to use the term intentional states for the latter meaning (and the more neutral term

Michael Tomasello, Max Planck Institute for Evolutionary Anthropology, Department of Developmental and Comparative Psychology, Germany, e-mail: tomas@eva.mpg.de

 ©2016, Michael Tomasello, published by De Gruyter.

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

psychological states when I want to talk about non-intentional states like perceptions and emotions).

These obvious rough spots aside, there was a consistent theme from all of the commentaries pressing on a substantive issue for which I have previously been criticized. The issue is that, in light of the neo-pragmatist and communitarian spirit of the approach in general, my treatment of young children's skills of joint intentionality at one and two years of age (as somewhat representative of early humans) is both too generous and too individualistic (neo-Cartesian). The even stronger criticism is that the way I characterize joint intentionality at one and two years of age already presupposes the outcome – propositional representations with “objective content”, to use Satne's characterization – that I am attempting to explain. But each of the commentators has a slightly different take on the issue, and so let me deal with them each in turn (touching on some other important issues in the process). I will say something about their alternative proposals after that.

Satne opines that in neo-pragmatist approaches in general there is an essential tension because people want to explain objective thinking (with objective propositional content) by virtue of participation in socio-cultural practices. But participation in these practices requires “intelligence”, which requires “intentionality”, which requires “objective content”. She applauds my two-step approach employing joint intentionality as a middle step (on the way to a full-blown collective intentionality characterized by objective thinking) as helpful in breaking down the problem; but then she claims that joint intentionality, as I described it, has essentially the same problem. Moll puts it this way: how can participation in collaborative activities lead to skills of joint intentionality, since participation in such activities requires skills of joint intentionality in the first place. “If a complex social structure with cooperative activities between hominins was already in place and sustained over extended periods of time, what purpose does a subsequent cognitive turn in the form of a ‘cooperativization’ serve?” Both Satne and Moll thus think that to begin to participate in human-like collaborative activities requires precisely the skills of joint intentionality we are attempting to explain.

But this is where the evolutionary methodology – as opposed to pure analysis – is critical because we can invoke a dialectic. Great apes are already “intelligent” and they can already engage in “collaborative” activities. The thinking underlying this collaboration is already abstract (schematizing non-propositional perceptual content) and its content is not punctate objects and events but rather situations or states of affairs (having the complexity, but not the perspectival [intensional] nature, of propositional content). And this is sufficient to engage in a certain form of collaboration, to understand others as intentional agents, etc. But then something happened. New circumstances arose such that, given

ape-like individuals who engaged in collaborative behavior based on individual intentionality, those who had cognitive skills that increased the effectiveness or efficiency of the behavior were at an adaptive advantage. Those individuals who, for reasons of random genetic assortment, were able to form joint goals (and peaceably share the spoils with others, etc.) proliferated at the expense of those who could not. And so, adapting Satne's formulation, we may say that participation in the practice at some initial level requires great ape intelligence and individual intentionality, but then the practice itself – in the context of new selective pressures – created a new adaptive context within which further skills of shared intentionality were naturally and socially selected.

In an analogous vein, Satne says in another place that the understanding of perspective – characteristic of joint intentionality – presupposes “objective content” – characteristic of later emerging collective intentionality – because it presupposes an objective referent of the different perspectives. But the point here is that if we think of “objectivity” as essentially intersubjectivity of a certain kind, then the notion of perspective only requires two individuals with differing perspectives on an object of joint attention. And then collective intentionality generates “objectivity” not by simply multiplying perspectives to the group at large, but, as I thought I said clearly, by making an in principle judgment: the perspective of anyone who would be one of us, any rational person, so that the objective perspective is in fact a perspectiveless view from nowhere. And so once again we need a dialectical process to go from a kind of dyadic perspective taking to a collective, or even “objective” perspective.

Koreñ has essentially the same problem, namely, that my description of joint intentionality presupposes the explanandum:

Tomasello's appeal to standard philosophical accounts misfires if his ambition is in part to show that and how prelinguistic abilities to engage in joint activities – present in children in their second year – could contribute to development of higher cognitive capacities such as full-blooded social-recursive mindreading. For standard accounts presuppose sophisticated cognitive skills of precisely this sort, so cannot shed light on their development via capacities for participating in joint action.

But here again I was probably not clear enough about my naturalistic mode of operation. I cite Bratman and others, but I often extract just one idea from the overall framework, and sometimes modify it slightly. So I did not say that I adopt Bratman's analysis in toto, but I stated my analysis and cited Bratman for the inspirational idea. In particular, I have used the term joint goal in preference to joint intention to avoid any confusion with intentionality as a broader characterization of thinking. Again drawing on the analysis of ape thinking, which includes individual attention, joint attention represents a joining of the attention of two

individuals, each of which is already structured by schematic representations of perceptual experience of whole situations. Joint intentionality is not propositional – if that requires objective content – but it has the seeds of this by joining together individual intentionalities that already have non-propositional perceptual content of fact-like situations. And so I am again attempting to invoke a kind of dialectic in which great apes' individual intentionalities (based on non-propositional content) are the starting point. Joint intentionality joins these together in a new way but without, as yet, any objective content; that is still to come. The problem no doubt – using the quote from Davidson that served as the epigraph of the final chapter – is that “[W]e lack ... a satisfactory vocabulary for describing the intermediate steps”.

Schmitz has a similar problem with my invocation of an “assumption of cooperation” (or “mutual assumption of cooperation”) necessary for cooperative communication, and here he is absolutely correct. Use of the term assumption was unthinking on my part; I had no clear idea of its cognitive status. But being pressed now, I guess I would simply see it as two individuals perceiving and understanding together – in joint attention – that each is trying to be cooperative. The evolutionary route to the shared understanding would once again be a dialectical process where initially it is each individual perceiving the other as a potential cooperator, but then each having a recursive understanding that the other sees them in the same way as well. Ultimately, in this domain and other domains involving joint attention and common ground, I continue to defend the view that the evolutionary process must have involved some kind of recursive embedding of psychological states. I labeled this as one of the deep outstanding problems which I know I have not solved, as Schmitz notes, but the evolutionary problem requires that we somehow get from individual intentionality to shared intentionality. It could have happened in one big leap from individual to shared, but it seems likely to me – especially given that we do have some capacities of recursive mindreading in certain situations, especially when our assumptions of sharedness break down – that an intermediate step to the phenomenological experience of sharing psychological states with others is a recursive step of the form: he sees me seeing the banana.

So what are the alternatives that the commentators propose? The term enactive was used by both Satne and Schmitz, whereas Koreň refers to minimalist accounts. My problem with enactive accounts is not that they are wrong, but that they are vague at all the key places. Schmitz's is as follows:

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

What is attentional structure? What are actional and perceptual sensibilities? What are hunches or instincts? Do sensory-motor-emotional schemata not have any cognitive structure? In a previous criticism of our approach Carpendale and colleagues (2013) say “the starting point is activity, and mental states and observable behavior are both aspects of activity” and this activity is sensitive to context”. But what on earth does it mean to be an aspect of activity or to be sensitive to context if one wants to define those terms without reference to any forms of cognitive structuring? I would very likely have no problem with Schmitz’s general characterization if he was forced to specify some of these terms in a more psychologically textured way. I understand that the enactive approach resists this move, but then they simply have pushed off the difficult problems into such things as perception, attention, and emotion – seemingly just to keep out cognition. I’m afraid, as a cognitive psychologist of a sort, I just do not see it. Schmitz’s idea of co-subject is congenial with my account, I think, as it represents another way of characterizing joint agency.

Koreň makes approving reference to three minimalist approaches that are not “enactive”, in the normal meaning of that term, but nevertheless eschew cognitive terminology, especially for one and two-year-old infants. First, he invokes Pacherie’s team reasoning approach in which individuals do not engage in complex inferences based on abstract representations, but “simply” understand themselves as part of the team and work for the good of the team. Perhaps I am being thick here, but that just seems to me to be talking metaphorically about joint intentionality. What does it mean to be a member of the team? Is this really easier to understand than a joint goal? Koreň also invokes Bermudez, but Bermudez basically just has a different approach for chimpanzees (perceptual mindreading). Koreň’s suggestion that something along these lines might fit prelinguistic infants is just not consistent with the data; prelinguistic infants engage with others in joint attention and cooperative communication in ways that great apes simply do not. Finally, Koreň also invokes Butterfill, characterizing his approach: “a shared goal is an outcome toward whose realization each participant directs her action, each participant having behavioral expectations to the effect that (a) other participants would perform an action direct at the goal and that (b) the outcome would be realized as a common result of goal-directed actions of all of them”. This is basically a characterization of chimpanzee group hunting of monkeys: each of them knows that each of them is individually trying to capture the monkey and that success will result only if they all try individually. I’m sorry,

but this is just way too thin for joint attention which require some kind of shared agency based on a common ground understanding of what are doing.

Finally, in a generally enactive spirit, Satne also has a somewhat different proposal. Her problems with the intermediate step of joint intentionality lead her to basically try to avoid it altogether. She proposes that we should just think about children becoming enculturated into the wider group from the beginning, without any intermediary stage in which they form special kinds of relationships and engage in special kinds of activities jointly with other individuals. There is no problem in principle with this approach, and I am sure it applies to some animal species like fish who just grow up in a group without forming any kinds of special relationships with individual others. And I am open to the idea that children are adapting to the group in parallel to their forming special kinds of individual relationships with others. But I simply do not see how we can leave out of account the kinds of second-personal relationships that are so important to human social interaction, for example, joint commitments that are made to other individuals. It also seems difficult to me to go straight to some kind of objective perspective without going through perspective taking of another individual first or to form collective goals with the group rather than joint goals with another individual. And I would say that empirically it is the case that infants and young children interact in quite sophisticated ways when interacting with a single other individual, and when they are in groups they mostly simply act in parallel with others or they engage with another individual. In any case, as I said, I accept the proposal – indeed, I think it has some merit – that joint and collective intentionality in some sense develop in parallel. But it is just an empirical fact that young children are first skillful with joint intentionality before they are skillful with collective intentionality, and it seems obvious that the kinds of relationships and interactions that children have experienced with other individuals should have some effect on the way they relate to and interact with the group.

Finally, I would like to respond to the very important point raised by Moll. She believes it is of utmost importance not to think of shared intentionality as an additional skill that humans have evolved, but rather as a transformative mode of operation. This is something with which I could not agree more. Classical approaches in evolutionary psychology, for example, those of Tooby and Cosmides, are modular and so additive. Each module has its own adaptive conditions and its own evolutionary function. But what I tried to argue in this book was that shared intentionality is no ordinary adaptation because it is an adaptation concerned with how one relates to others and this affects all of the relationships and interactions that have a social dimension, including interactions with the physical world in so far as others – or their products, such as a language – are involved. A central point of the book, in addition, is that these forms of interaction change

humans' modes of cognitive representation, inference, and self-regulation. And so indeed, to quote myself, the emergence of shared intentionality “changed everything”.

Again let me thank my commentators, along with Hans Bernard Schmid who organized the symposium. For many years I searched for theoretical help with my empirical problems in psychology and cognitive science – in vain. I discovered that it is only in philosophy that people take seriously the social and cultural dimensions of human cognition and thinking, including its normative structuring. Having borrowed from philosophical theories of shared intentionality, I have attempted to give back a bit as well, though I understand that the use of philosophical concepts (sometimes uprooted from their original soil) to explain empirical phenomena may or may not actually make a philosophical contribution.

Bibliography

- Carpendale, J., S. Atwood and V. Kettner (2013): “Meaning and Mind from the Perspective of Dualist versus Relational Worldviews”. In: *Human Development* 56, p. 381–400.
- Tomasello, M., M. Carpenter, J. Call, T. Behne and H. Moll (2005): “Understanding and Sharing Intentions: The Origins of Cultural Cognition”. In: *Behavioral and Brain Sciences* 28, p. 675–691.