

Book Symposium

Open Access

Ladislav Koreň

Joint Intentionality

From Thin to Thick

DOI 10.1515/jso-2015-0047


Abstract: According to the *shared intentionality hypothesis* proposed by Michael Tomasello, two cognitive upgrades – *joint* and *collective intentionality*, respectively – make human thinking unique. Joint intentionality, in particular, is a mindset supposed to account for our early, species-specific capacity to participate in collaborative activities involving two (or a few) agents. In order to elucidate such activities and their proximate cognitive-motivational mechanism, Tomasello draws on philosophical accounts of shared intentionality. I argue that his deference to such cognitively demanding accounts of shared intentional activities is problematic if his theoretical ambition is in part to show that and how early (prelinguistic and precultural) capacities for joint action contribute to the development of higher cognitive capacities.

Keywords: Tomasello; Cooperation; Joint intentionality; Joint action; Mindreading.

1 Introduction

Michael Tomasello's (2014) much anticipated book, *A Natural History of Human Thinking*, provides a synthesis of his research to date on the cognitive infrastructure of human cooperative lifeways. According to his *shared intentionality hypothesis*, two upgrades on cognitive skills of apes make human cognition unique – *joint* and *collective intentionality* respectively. Joint intentionality, in particular, is a mindset supposed to account for our early, species-specific capacity to participate in collaborative activities involving two (or a few) agents. In order to elucidate such activities and their proximate cognitive-motivational mechanism,

Ladislav Koreň, Institute of Philosophy, Academy of Sciences of the Czech Republic, Jilská 1, Prague 110 00, Czech Republic, e-mail: lacokoren@gmail.com

 ©2016, Ladislav Koreň, published by De Gruyter.

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

Tomasello draws on philosophical accounts of shared intentionality (Gilbert 1989, 1990; Bratman 1992, 1993; Searle 1995; Tuomela 2007). I argue that his deference to such cognitively demanding accounts of shared intentional activities is problematic if his theoretical ambition is in part to show that and how early (prelinguistic and precultural) capacities for joint action contribute to the development of higher cognitive capacities.

2 The Shared Intentionality Hypothesis

What does it take to *think*? According to Tomasello, a thinking organism is a self-regulating system capable of processing, storing and evaluating information from its environment and utilizing it to realize its reference goals by flexibly adjusting its behavioral strategies to (often novel) situations occurring in its dynamically changing habitat (p. 7–9). Furthermore, truly characteristic of thinking is the capacity to simulate and evaluate such experiences and processes in “offline” regime: imagining and evaluating, ahead of time, alternative behavioral strategies of realizing a reference goal in given circumstances. So characterized, thinking builds on three core cognitive capacities: (1) *representing* relevant features of situations in a schematic-generalized form of a sort; (2) connecting such representations in *inferences* that transform them according to causal, intentional or (proto-)logical relations; and (3) *self-monitoring* goal-directed behavior by controlling its execution and by evaluating its success with respect to goal-attainment.

Tomasello is ready to attribute the core cognitive skills also to non-human animals, and to primates in particular. Indeed, assuming some basic cognitive (as well as motivational) commonalities with our nearest evolutionary relatives – the great apes – he asks: What cognitive (and motivational) updates on ape-like cognitive skills (motivations) can be isolated as human-unique innovations, and what evolutionary pressures could have been selected for them? He argues that the main difference does not lie in our cognitive ways of coping with physical domains but in our special ways of coping with ever more complex social domains.

On the one hand, Tomasello claims that apes (certainly great apes) share with humans cognitive skills (and motivations) enabling them to “read” intentional behavior of conspecifics by means of representing, monitoring and inferring goals, perceptions and knowledge of one another. Against behavioral accounts of primate social cognition,¹ he argues that apes possess impressive abilities

¹ See Penn and Povinelli (2007) who argue that available evidence is compatible with a more parsimonious hypothesis according to which apes are *behaviour*-readers rather than *mind*-readers.

to think of one another not just as animate agents but also as agents causally manipulating the environment in order to realize their reference goals, with a multitude of behavioral opportunities and a cognitive perspective (what they can perceive or know) bearing on the prospects of goal-attainment. In particular, they are capable of causally, intentionally, even proto-logically transforming representations of situations occurring in their socio-physical habitat so as to infer how conspecifics might behave in them, based on what conspecifics are represented as *perceiving/knowing* (or *not perceiving/knowing*) and on the goals they are likely to pursue.

Now, because only behavior is observable, the question arises what warrants attribution of *mindreading* skills to apes allegedly enabling them to represent mental-cognitive states (perceptual states, knowledge) as causes of behavior and to inferentially process representations of such states to anticipate/predict behavior of conspecifics. Here, Tomasello submits, cleverly devised experimental measures look promising, such as famous food-competition paradigms of Hare et al. (2000, 2001). In Hare et al. paradigms experimenters had a dominant (X) and sub-ordinate chimpanzee (Y) compete over food (F), where some pieces of food were visible to both Xs and Ys and some only to Ys. It turned out that Ys more often pursued F that Xs were not in a position to see. So Ys displayed a tendency to adjust their behavioral responses as if anticipating Xs' behavior based on representing Xs' cognitive position vis-à-vis F. Tomasello's conjecture is that there is an understanding on the part of Ys that Xs will pursue F (which Ys can see) only if Xs see F (or if Xs know where F is, because Xs saw F there a while before). If, then, Ys see (know) that Xs are in a position to see F (or that Xs are in a position to know where F is), Ys might infer that Xs will pursue F, in which case Ys would not risk pursuing F. If, on the other hand, Ys see (or know) that Xs are not in a position to see F (or that Xs do not know where F is), Ys might infer that Xs won't pursue F, and hence Ys might decide to pursue F. This explanation of observed patterns of behavior attributes to chimpanzees a mindreading mechanism of a sort, allowing them to represent and infer mental-cognitive states.²

On the other hand, Tomasello contends that conspicuously absent from the repertoire of goal-directed activities of apes are joint collaborative activities ubiquitous in human social interactions, in which already human infants engage. He reviews a wealth of telling studies of infants (many in their second year, or even earlier) documenting their developing abilities to coordinate their attention with others and engage in joint (playful) activities, willingness

² If one that falls short of a full-blown mindreading capacity involving representations of belief-type mental states – the litmus test being success on a variety of false-belief tasks. Cf. Call and Tomasello (2008).

to help others irrespectively of external rewards (e.g. by handing or otherwise manipulating relevant objects) or to share information bearing on the task at hand (altruistic pointing). None of this, he suggests, has been documented in comparative studies of apes [except for interesting but rather limited cases of helping exhibited by human-reared chimps (Warneken et al. 2006)]. While apes, like humans, can think and reason about one another as “having” (and acting in pursuit of) goals – along with experiences, perceptions and knowledge bearing on goal-directed behavior – they seem incapable of “sharing” attention, experiences, perceptions or knowledge, hence of forming shared goals and engaging in genuinely shared intentional activities.

What accounts for this striking difference? The key to the answer is *cooperation*: “Great apes are all about cognition for competition. Human beings, by contrast, are all about (or mostly about) cooperation.” (p. 31). Whereas the *Machiavellian intelligence hypothesis* might be on the right track when it comes to explaining evolutionary origins of ape-style egocentric thinking (Byrne and Whiten 1988), Tomasello’s *shared intentionality hypothesis* aims to explain precisely why humans differ dramatically from their nearest relatives, being predisposed to acquire, early in ontogenesis, skills to experience, think, reason and act *together* that enable pursuit of shared goals and plans in accordance with division of roles and responsibilities (including commitment to realize a shared intentional activity and assist one another). These abilities (supported by requisite motivations towards greater tolerance, sharing, helping, etc.) are human innovations and Tomasello speculates that they could have developed (hand in hand with rudimentary skills of cooperative communication) once early (prelinguistic and precultural) humans were pressed to adapt to foraging niches in which cooperation-for-collaboration, rather than cooperation-for-competition, was a key to securing higher fitness. Tomasello calls them skills of *joint intentionality* underlying joint collaborative activities structured around shared goals and plans/intentions. Drawing on Bratman’s (1992) account of *shared cooperative activity*, he explains that in joint intentional activities interactants mutually recognize – in the situational common ground that can, up to a point, be made explicit in a mindreading recursion – the dual structure of shared goal and different individual roles, coordinating on this basis their role-based activities (meshing their sub-plans) and assisting one another in performing them (e.g. by sharing information) (p. 38, 40). In Tomasello’s larger picture, such skills form a prerequisite for all major cognitive upgrades to come – linguistic competence included – making human infants prepared to gradually absorb ever more sophisticated socio-cultural skills not only via observation and imitation of relevant others but also via active cultural pedagogy by adult teachers. This makes them prepared to acquire higher-level human-unique skills of *collective*

intentionality involving an ability to adopt a perspective of a group-member, acting as its deputy in both conforming to and enforcing its social norms, etc.

3 Discussion

Tomasello's shared intentionality hypothesis derives much of its attraction from the fact that it approaches human thinking from the developmental perspective, aiming to illuminate how higher cognitive skills of representation, inference and self-monitoring build on more rudimentary skills. The process is reconstructed as one in which human cognitive and communicative skills gradually evolve into ever more cooperative forms, the two milestones being the development of joint and collective intentionality respectively (in this order). From this developmental perspective, a pertinent question to ask is what motivations and cognitive abilities can be isolated as underlying early forms of joint activities – viz. those in which prelinguistic infants start to engage in their second year – and how they contribute to the later development of higher socio-cognitive skills that (together with appropriate motivations) provide cognitive infrastructure for ever more complex forms of collective actions. For instance, evidence from various developmental studies indicates that it takes quite some time for children to acquire full-blown mind reading skills enabling them to represent (as well as distinguish from one another) propositional mental states – paradigmatically, *belief*, *desire*, *intention* – with full-blown understanding of their specific conditions of success.³ Among other things, Tomasello's developmental (and comparative) approach promises to shed important new light on this process precisely by focusing on early forms of human joint activities and isolating their proximate cognitive-motivational mechanisms.

So conceived, however, the developmental approach calls for an account of early joint activities – and their cognitive-motivational infrastructure – that does not presuppose what it aims to illuminate, namely higher cognitive skills for representation and monitoring of propositional mental states and inferences operating on such *meta*-representations (Butterfill 2012). Now, Tomasello draws inspiration from standard philosophical accounts of shared intentionality (henceforth, SPAs) that provide analyses of what it takes to do something together based on sharing goals and intentions (Gilbert 1989, 1990; Bratman 1992, 1993; Searle 1995; Tuomela 2007). But there is an immediate problem with invoking SPAs to shed light on the

³ It is often considered evidence of this gradual development of mindreading skills in children that they do not reliably pass a variety of false-belief tasks until late in their fourth year.

nature of early joint activities and their cognitive underpinnings. First, SPAs have been designed to fit paradigmatic cases of shared activities of mature/socialized human agents – e.g. *walking together* (Gilbert 1990), *preparing a food together* (Searle 1990), *singing a duet together* (Bratman 1992), etc. Second, though they imply different claims about the nature of shared intentions and actions based on them, SPAs tend to presume demanding cognitive machinery behind shared intentions and activities.

Consider the following formulation of Bratman's account of shared intention – one required for a shared cooperative activity J – to which Tomasello refers:⁴

1. (a) I intend that we J and (b) you intend that we J.
2. I intend that we J in accordance with and because of la, lb, and meshing subplans of la and lb; you intend that we J in accordance with and because of la, lb, and meshing subplans of la and lb.
3. 1 and 2 are common knowledge between us. (Bratman 1993, p. 106)

For agents to “share intention” in this sense – and to engage in a shared intentional activity presupposing this kind of intention sharing – they must not only (first-order) represent non-mental states-of-affairs but also (second-order) represent propositional mental states (viz. intention, knowledge) of themselves/others, and to iteratively embed such meta-representations in one another.⁵ But this would seem to require a significant cognitive sophistication (mindreading skills) on the part of agents participating in shared collaborative activities of the sort that the developmental approach aims to elucidate. Let me expand on this point a bit.

In his two-stage reconstruction of the evolution of human thinking, Tomasello concedes that skills of joint intentionality developed by ancestral humans before the advent of collective intentionality and conventional linguistic representations did not yet make available to them the concept of objectively true (or false) representation (p. 86–87). Hence it is problematic to ascribe to prelinguistic humans the concept of *belief*, which presupposes the idea of a representation being objectively true or false. Tomasello concurs, arguing that *I-Thou* relations involved in triangulation do not yet provide any idea of objective representation, hence the concept of propositional belief. And he concludes (exploiting an ontogeny-phylogeny parallel):

⁴ Compare this: “We may characterize the formation of a joint goal (or joint intention) in more detail as follows (see Bratman 1992). For you and me to form a joint goal (or joint intention) to pursue a stag together, (1) I must have the goal to capture the stag to get her with you; (2) you must have the goal to capture the stag to get her with me; and, critically, (3) we must have mutual knowledge, or common ground, that we both know each other's goal.” (p. 38)

⁵ This has been pointed out by several writers. Cf. Butterfill 2012 and Pacherie 2011.

It is likely that young children begin to think in terms of multiple different perspectives on things from as soon as they participate in joint attention with its two perspectives during late infancy [...] and we may hypothesize that this was the case for early humans as well. But it is not for several more years that children come to a full-blown understanding of beliefs, including false beliefs, because they (and so presumably all humans before modern humans) do not yet understand “objective reality”. (p. 87)

So, while Tomasello is adamant that young children in their second year participate in joint activities displaying rudimentary skills of joint attention and intentionality, he admits that the notion of belief might be available to children no sooner than around their fourth year (so when they already speak and can track false beliefs, as evidenced by the fact that they to pass a variety of false-belief tasks). What is more, Tomasello and Rakoczy write in their joint paper that: “Following a growing number of researchers, we believe that a critical role in children’s construction of a belief-desire psychology – understanding persons as mental agents – is played by processes of linguistic communication.” (Tomasello and Rakoczy 2003, p. 134) They continue with providing a sketchy account of how children could gradually acquire the concept of belief due in part to learning to master sentential complement constructions. In *A Natural History of Human Thinking*, I take it, Tomasello proposes an analogous conjecture about the phylogeny of human cooperative thinking (p. 103–104): higher cognitive skills for ever more complex cooperative activities – including full-blown understanding of beliefs, hence full-blown mind-reading skills – co-developed with fully conventional linguistic communication as a successor of cooperative pointing and pantomiming.

But if we grant that prelinguistic (and precultural) creatures (infants or our pre-modern ancestors) do not yet display a full-blown understanding of beliefs, what about their understanding of knowledge, desire or intention? Philosophers, including proponents of SPAs, standardly assume that, like belief, knowledge, desire and intention are (a) mental states individuated by their propositional contents, (b) interlinked in various ways reflecting their causal-explanatory role, (c) displaying intensional features, and (d) being subject to certain rational constraints (e.g. of practical reasoning). Thus, in our folk psychology, *intention* construed as an action-plan to realize a certain goal seems closely tied to a *belief* to the effect that certain means are suitable to realize the goal. In addition, there are rational constraints on intentions such as consistency with an agent’s beliefs as well as with other intentions of an agent. Given this link between intention and belief, prelinguistic and pre-cultural creatures may well lack a full-blown understanding of intentions so conceived. In a similar vein, traditional philosophical consensus about *knowledge* (hence about common knowledge) is that knowledge involves *true belief* (or a belief-type state) as a necessary ingredient. Once we acknowledge this link between knowledge and belief, prelinguistic

and precultural creatures would seem to lack also a full-blown understanding of knowledge so conceived, because a full-blown understanding would require them to represent knowers as (true) believers.⁶

Summing up, the worry raised by previous considerations is that if Tomasello is to keep his developmental approach, SPAs might not be his most natural allies in illuminating the nature and cognitive-motivational infrastructure of early forms of joint activities. Assuming that the worry is well taken, we are left with the question: What kind of account, assuming what kind of mindreading skills, would better fit joint activities of prelinguistic creatures?

Some theorists have recently urged us to look for *modest* or *minimal* accounts that would have us assume less by way of cognitive demands on coordination – alignment and adjustment – of attention, experience and action. For example, Bermúdez (2009, 2011) motivates a distinction between *perceptual* and *propositional* mindreading. The former requires only representation of perceptual (registering) relations between subject and object/state of affairs (which might be indicated to a mindreader, say, by the subject's direction of gaze), but the latter requires understanding of a more complex logical-inferential-rational space of finely articulated propositional contents that, Bermúdez argues, opens up only with mastery of natural language.⁷ Bermúdez then makes a compelling case that,

⁶ But is it not open to Tomasello to claim that a variety of knowledge (or, perhaps, proto-knowledge) may be represented without representing belief-type mental states? Thus, given that knowledge is factive, it may be suggested that it requires less cognitive sophistication to represent X as knowing *what is the case* (where the question *Truly or falsely?* does not arise) than to represent X as believing something (where the question does arise). I do not mean to deny that one may talk of *proto-knowledge* in some such sense (or of proto-intention in some related sense). But then I want to point out that there is another aspect of knowledge emphasized by philosophers who think that knowledge is closely tied to belief-type states, which is that conceptualizations in terms of *know-that* constructions display intensional features such as referential opacity, as evidenced by the fact that substitution of co-referential terms within the scope of complements might change the truth-value of knowledge ascription. To the extent that it is plausible to claim that “a critical role” in the development of full-blown understanding of propositional mental states displaying *intensional* features is played by processes of “linguistic communication”, it is not clear what would warrant attribution of such an understanding to pre-linguistic and precultural creatures. I am grateful to an anonymous referee for pressing me to make this point clearer.

⁷ The point is that a mindreader X that utilizes propositional attitude attributions to infer the behaviour of another agent Y would typically need (a) to represent explicitly Y's background psychological profile (including further propositional states) and (b) to simulate to some extent Y's practical reasoning coherently relating the attitude to the background profile. This seems to require that finely articulated representations of propositional states be available to X in a conscious format, which only a medium such as natural language seems capable of providing. By contrast, the relation between a perceptual state of Y and her behaviour can be represented by X in a more direct way, since X would typically perceptually register the state of affairs that

up to a point at least, perceptual mindreading can account for experimental evidence of ape understanding of goal-directed agency of conspecifics [including, in particular, the object choice and food-competition paradigms employed by Tomasello and his associates (Hare et al. 2000; Hare et al. 2001; Hare and Tomasello 2004)] without imputing to mindreaders any capacity to represent and recursively reason about propositional states of themselves and others (e.g. what other chimps know or are ignorant of).

One wonders whether something along similar lines could not fit also *some* cases of prelinguistic children – or, for that matter, of early humans before the advent of language and culture – participating in collaborative activities. Indeed, pursuing a somewhat similar line as Bermúdez, Butterfill and Apperly (2013) have developed a *minimal theory of mind*, according to which a creature X equipped with rather modest mindreading skills could *track* propositional mental states of Y not by directly representing them as such, but rather by representing their proxies such as Y's *goals* (outcomes to be realized as a function of Y's behavior), *encounterings* (of Y with objects in Y's vicinity) and *registrations* (correct or incorrect, depending on whether an object recently encountered by Y at a location L is or is not at L).⁸ This theory can not only account for observed patterns of chimpanzee behavior in food-competition paradigms but may have some promise when it comes to elucidating early forms of human joint activities. In an earlier article Butterfill (2012) argues that, in place of shared intention *à la* Bratman, there is room – and developmental/evolutionary rationale – for more modest accounts of joint action, capitalizing on the notion of a *shared goal* around which a plural activity of multiple agents is coordinated in the sense that: 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 directed to the goal and that (b) the outcome would be realized as a common effect of goal-directed actions of all of them.⁹ This account is compatible with the minimal theory of mind of Butterfill and Apperly (representations of goals, encounterings and registrations on the part of interactants might be vital to early joint activities), requiring neither full-blown understanding of knowledge, beliefs or intentions (hence no higher-order intentions implied in Bratman-style account of shared intentional activities), nor, for that matter, common knowledge on the part of agents coordinating their actions around a shared goal.

Y registers and could draw on her own previous experiences of a like sort (viz. links between perceptions and behaviour). In this way, the cognitive load of perceptual mindreading is less heavy than that of propositional mindreading.

⁸ For precise accounts of these notions the reader should consult Butterfill and Apperly 2013.

⁹ This is a simplified version of Butterfill 2012, p. 40–43.

Admittedly, there are more complex forms of joint action sustained by cognitively more demanding mechanisms, up to those that involve full-blown mindreading and second-personal self-monitoring. Still, minimalist accounts of joint-coordinated activities seem particularly qualified to play a vital explanatory role in the developmental approach to specifically human social cognition. From this perspective, some bootstrapping approach might be called for, discriminating various levels of joint action (shared goals, intentions), ranging from cognitively modest proto-versions to ever more complex and cognitively demanding ones.

Acknowledgments: My work on this study was supported by the grant GA13-20785S, *The nature of the normative – ontology, semantics, logic*. I am grateful to an anonymous referee for helpful comments.

Bibliography

- Bermúdez, J. L. (2009): “Mindreading in the Animal Kingdom”. In: R. W. Lurz (Ed.): *The Philosophy of Animal Minds*, New York: Cambridge University Press, p. 145–164.
- Bermúdez, J. L. (2011): “The Force-Field Puzzle and Mindreading in Non-Human Primates”. In: *Review of Philosophy and Psychology* 2. No. 3, p. 397–410.
- Bratman, M. E. (1992): “Shared Cooperative Activity”. In: *Philosophical Review* 101. No. 2, p. 327–341.
- Bratman, M. E. (1993): “Shared Intention”. In: *Ethics* 104. No.1, p. 97–113.
- Butterfill, S. (2012): “Joint Action and Development”. In: *Philosophical Quarterly* 62. No. 246, p. 23–47.
- Butterfill, S. A. and I. A. Apperly (2013): “How to Construct a Minimal Theory of Mind”. In: *Mind & Language* 28. No. 5, p. 606–637.
- Byrne, R. W. and A. Whiten (1988): *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans*. New York: Oxford University Press.
- Call, J. and M. Tomasello (2008): “Does the Chimpanzee Have a Theory of Mind? 30 Years Later”. In: *Trends in cognitive sciences* 12. No 5, p. 187–192.
- Gilbert, M. (1989): *On Social Facts*. London: Routledge.
- Gilbert, M. (1990): “Walking Together: A Paradigmatic Social Phenomenon”. In: *Midwest Studies in Philosophy* 15. No. 1, p. 1–14.
- Hare, B., J. Call, B. Agnetta, and M. Tomasello (2000): “Chimpanzees Know What Conspecifics Do and Do Not See”. *Animal Behaviour* 59. No. 4, p. 771–785.
- Hare, B., J. Call, and M. Tomasello (2001): “Do Chimpanzees Know What Conspecifics Know?” In: *Animal Behaviour* 61. No. 1, p. 139–151.
- Hare, B. and M. Tomasello (2004): “Chimpanzees are More Skillful in Competitive than in Cooperative Cognitive Tasks”. *Animal Behaviour* 68. No. 3, p. 571–581.
- Pacherie, E. (2011): “Framing Joint Action”. In: *Review of Philosophy and Psychology* 2. No. 2, p. 173–192.

- Penn, D. C. and D. J. Povinelli (2007): "On the Lack of Evidence that Non-Human Animals Possess Anything Remotely Resembling a 'Theory of Mind' ". In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 362. No. 1480, p. 731–744.
- Searle, J. (1990): "Collective Intentions and Actions". In: Philip R. Cohen, Jerry Morgan, and Martha E. Pollock (Eds.): *Intentions in Communication*. Cambridge Mass.: Bradford Books, MIT Press.
- Searle, J. (1995): *The Construction of Social Reality*. New York: Free Press.
- Tomasello, M. and H. Rakoczy (2003): "What makes Human Cognition Unique? From Individual to Shared to Collective Intentionality". *Mind & Language* 18. No. 2, p. 121–147.
- Tomasello, M. (2014): *A Natural History of Human Thinking*. Cambridge MA: Harvard University Press.
- Tuomela, R. (2007): *The Philosophy of Sociality: The Shared Point of View*. New York: Oxford University Press.
- Warneken, F., F. Chen, and M. Tomasello (2006): "Cooperative Activities in Young Children and Chimpanzees". In: *Child Development* 77. No. 3, p. 640–663.