

Joint action and teleological action understanding

Denis Tatone – January 07, 2013

The many routes to joint action

As Butterfill notes (2012), the topic of joint action can be –and has traditionally been – approached from different perspectives: psychologically, to specify which mechanisms make it possible to understand and engage in different sorts of joint action; developmentally, to analyze which socio-cognitive abilities it presupposes and which it may scaffold; conceptually, to provide a principled way of distinguishing joint action and joint reasoning from their individual counterparts; phenomenologically, to characterize how the experience of shared agency differs from cases of shallow interaction between agents; metaphysically, to define what kind of entities or attitudes are implied by the recognition that some actions are joint; and normatively, to spell out what kind of commitments (if any) are imposed on the participants of joint action.

Each of these dimensions has been emphasized to various extents by different scholars. Gilbert (1992), for example, argued that the distinctiveness of joint action has to be found in the special kind of (non-moral) commitments and obligations to others it require, which re-qualify the authors of such commitments as “plural subjects” of the joint action. In her account, the normative component of joint action is thus crucial to constitute a new metaphysical entity, the plural subject, which is supposed to capture the peculiarity of the phenomenon. Opposing to this and other views, in sketching our account of joint action we intend to closely follow Bratman’s theoretical guidelines to “not appeal to an attitude in the mind of some superagent, nor [...] assume that shared intentions are always grounded in prior promises” (Bratman, 1993). As Pacherie (2011) puts it, the construction materials of Bratman’s account “come cheap” because it does not require novel kinds of agents, commitments or attitudes. However, their assemblage is costly and demands cognitively sophisticated agents. In fact, Bratman characterizes joint action – otherwise called “shared intentional activity” (Bratman, 1992) – in terms of shared intentions, which require forming (a) intentions about intentions, (b) intentions about meshing subplans of with the other agent engaged in joint action, and (c) the belief that the knowledge of (a) and (b) is common between the agents. Because of the sophisticated kind of propositional attitude ascription involved, Bratman’s account seems ill-suited to provide a plausible description of how agents lacking full-fledged metarepresentational abilities, such as animals and young children, can engage in and understand joint activities.

However, Moll & Tomasello (2007) and, more explicitly, Carpenter (2009) claim that children’s participation in joint action fully satisfies Bratman’s requirement, because toddlers exhibit understanding of common knowledge (from 14 months of age: Liebal et al., 2009), intention to engage in joint action with an adult by meshing subplans (from 18 months of age: Warneken et al., 2006), and commitment to the joint activity (robustly manifested from the third year of age: Gräfenhaim et al., 2009). We do not intend to discuss here how convincing are the evidences these studies provide in favor of Carpenter’s claim. However, even taking her conclusions at face value, Carpenter hypothesizes that the engagement in joint activities – in the sense articulated by Bratman – fosters the development of mindreading. As highlighted by Butterfill (2012), this argument contains an obvious contradiction: given that sophisticated mindreading is required for sharing intention, how is it possible that meeting the sufficient conditions for joint action could possibly explain the development of what it actually presupposes?

Setting the stage of joint action for mind-blind creatures

To avoid similar loopholes, we need an account of joint action that, while being cognitively and normatively parsimonious like Bratman's account, avoids its cognitive prodigality. Furthermore, in order to provide a characterization of joint action that best resembles the everyday shared activities children engage in, we will restrict our discussion to cases of "modest sociality" – that is, cases of small-scale shared intentional agency in absence of asymmetric authority relations (Bratman, 2009) – which (a) take place in "ephemeral groups" (i.e., groups whose identity as a group consists just in the fact that a set of person is acting jointly: Kutz, 2000) and in absence of shared preference and habits that can otherwise provide alternative ways to ensure coordination (Butterfill, 2012), and (b) are composed of "cooperatively neutral" act-types (Bratman, 1993), since any commitment to mutual helping is supposed to be explained by, and not be already pre-existent to, the shared activity. Moreover, the cases of joint action we aim to explain are those where it is discernible from an observer's (or an agent's) viewpoint that the coordination among agents' action, under normal circumstances, plays a role in bringing about the collective goal. The emphasis on the recognition of joint action via (visual) analysis of the actions' structure qualifies our account defined as "anti-internalist" in the sense that, not relying on unobservable binding entities (such as promises or we-intentions), depends crucially on temporal or spatial signatures of coordination for the joint action to be identified as such. This makes it difficult, if not impossible, for an observer/participant to identify unambiguously instances of joint action, such as walking together as described by Gilbert (1990), in which the loose topology of the task does not provide observable cues to distinguish cases of "authentic" joint walking from cases where each agent individually pursues the collective goal. Said otherwise, our account will be limited to situations where it can be specified in (at least) one way a specific temporal and spatial patterning of the action sequence that reliably "afford" an interpretation of these agents' actions as contributory relations to a collective goal – which may take instrumental form, if such actions do help cause the collective outcome (e.g., removing a tree blocking the road: Butterfill, 2012) or if they represent constitutive elements of the group act (e.g., stepping this way is part of dancing a tango: see Kutz, 2000).

Having restricted our discussion to a subset of cases of joint action, and having already made clear that our account assumes a deep continuity (conceptual, metaphysical, and normative) between individual and joint action, without requiring – unlike Bratman – sophisticated kinds of mindreading, we will proceed as follows. First, we will briefly review other accounts which do not require representing another agent's beliefs and intentions (or even cues which stand as proxies for intentions: Tollefsen, 2005). After that, we will sketch our theory in close comparison with the shared-goal account formulated by Butterfill (2011; 2012). Lastly, we will go back to the developmental playground to evaluate how our theory stands in relation to some recent findings about children's engagement (and understanding of) in joint action.

The risky business of trusting frames

One account considered able to capture "a form of joint action more substantial than what the minimalist approach is content with" while reducing the cognitive costs of shared intentions (Pacherie, 2011) is the team-reasoning theory articulated by Bacharach. As he puts it: "somebody team-reasons if she works out the best possible feasible combination of actions for all the members of her team, then does her part in it" (Bacharach, 2006). For an agent to engage in this form of team-reasoning, she must have a self-conception as a member of the team. This collective self-framing in turn depends on the priming of a "we-frame" able to produce what Bacharach calls a "participatory intention". In a game-theoretic scenario, this would correspond to a situation where two agents facing a game, such as the Prisoner's Dilemma, characterized by common interests (i.e., interests of both agents which, given two options, are better served in one

option than in the other), and strong interdependence (i.e., when there is an outcome of common interest which can only be achieved together and which is not preferred by both agents to all other feasible) make cooperative choices. As Pacherie notes (2011), this approach allows us to reduce the cognitive demands we impose on the agents, since all it is required to them is to conceive themselves as members of a group. This (even unconscious) collective self-framing does the trick of lead people to spontaneously team-reason, thus not even requiring them to consider whether others will be so inclined. Differently from Sugden's theory of team-reasoning, in which all members of the group are required to have a common reason to believe that each member endorse and acts on team reasons (Sugden, 2003), in Bacharach's version of team-agency theory assurance plays no role. Contrarily to Pacherie (2011), we think that this simplification of the team-reasoning theory pays has unwelcomed side-effects for the theory. In fact, while reducing its normative import, this full reliance on the "we-frame" to solve the problem of identifying team members makes the decision to cooperate exceptionally risky – in the sense that no evidence to justify the decision to engage in a potentially costly activity can be gathered by the players. In fact, given that a group payoff-maximization strategy depends on both agents choosing to cooperate, how can Player 1 be sure that Player 2 will engage in team-reasoning, if there is no common knowledge about whether both players have been primed with the collective frame? In light of this, it is rather curious that Pacherie sees as one of the main advantages of this account the fact that it allows cooperation to emerge in situations where agents cannot communicate (2011)! However, our aim is not to discuss the limitations of Bacharach's account per se, but to provide a critical argument against the possibility of articulating a way to produce joint action (in this case through a specific form of reasoning) without considering the conditions of its success. Here the unpredictability of the framing effect and the ignorance about the other player's self-conception makes the coordination attempts so exposed to failure to cast doubts on the plausibility of the mechanism envisaged.

Unpacking the identification problem in a perceptually overt world

Given how pressing is the identification of the agents participating in the joint action, is there an alternative to Sugden's common-knowledge requirement which does not require belief ascription? One is provided by Tollefsen (2005) who argues that the coordination that takes place in joint action can be accomplished via joint attention – that is, "by means that the other participant has a focus of attention to the same entity as the self". This does not require anything more than an understanding of the other participant "as a locus from which action can originate" – a cognitive feat which does not need a robust theory of mind. Minimally reconsidered, the requirement of common knowledge becomes one of mutual awareness, where virtually nothing of the other player has to be represented for joint action to happen (Vesper et al., 2010). This bare-to-the-bone requirement can be found in different flavors in Kutz (2010) and Butterfill (the second key feature of plural activity: 2012), where any agent able to engage in shared goal-directed actions qualifies as putative partner. Moreover, the condition of "perceptual overtness" (Tollefsen, 2005) which characterizes the everyday cases of joint action makes mutually manifest to the interactants the space of possibilities in which the action can take place. This makes possible for the agents, in a sense, to "communicate without communicating". Said otherwise, provided that the putative interactants are capable of teleological action understanding (more on this later), an agent that is performing an action (for example, grasping one of the handle of a treasure chest) whose execution satisfies certain efficiency parameters that qualify it as a means action to a shared goal (more on this later) is at the same time is providing a "diagnostic cue" of her likely engagement in the joint activity (cf. with the concept of "coordination signal" in Vesper et al., 2010). As it is clear, this process of identification of the other agent as participant in the shared activity depends crucially on observing her actions and on inferring from it either the goals they are directed to or – to use the more mentalistic account given by Tollefsen (2005) – the intentions behind it, which are "not formed in advance of the action, but rather [that] causes the act by representing its conditions of satisfaction on the

fly". This "externalist constraint" we impose on our account serves the purpose of restricting the range of potential solutions to the identification dilemma (i.e., how to recognize whether the other agent is likely to act jointly with me) to the interpretation of actions in conditions of a perceptually overt common space and no communication. While this constraint forces us to exclude internalistic accounts such as Kutz's (2000), in which joint action can be achieved by satisfying the condition of having a participatory intention (i.e., an intention to do my part of a collective act), at the same time – we believe – it grounds our investigation of joint action on ecologically valid cases, which mirror the statistically most recurrent conditions in which the scaffolding of joint action during development takes place (see: Brownell, 2011).

Importing teleological action understanding into the shared-goal account

Before introducing our theory, it is worth reviewing another account operating under similar (cognitive and situational) restraints. In two papers, Butterfill (2011; 2012) articulates a minimal account of joint action which keeps Bratman's normative and metaphysical "conservatism" (2009) – to be understood in the sense that (a) social norms emerge from the norms of individual planning agency, and (b) causation is not distinct from the underlying causal processes involving organized elements of the group) without the mindreading requirement. According to Butterfill (2012), what makes coordination of activities possible is the presence of a shared goal, identifiable by a number of key features: (a) "there is a single goal G to which each agent's actions are, or will be, individually directed"; (b) "each agent can identify each of the other agents in a way that doesn't depend on the knowledge of the goal or actions directed to it"; (c) "each agent expects each of the other agents she can identify to perform an action directed to the goal"; and (d) "each agent expects this goal to occur as a common effect of all of their actions directed to the goal". In the essence, Butterfill's claim is that "an expectation, on the part of each agent concerned, that she is or will be in a plural activity [i.e., an activity for which each agent's activities are individually organized around a single outcome which occurs as a common effect of all the agents' activities] with the others, will, in favorable circumstances, and *in concert with emergent coordination*, normally enable them to coordinate their actions" (my italics).

Our account is meant to be formulated along the same lines of Butterfill's. As we already wrote before, we too defined as pre-requisite to joint action (a) the mutual identification of agents and (b) the application of reasoning about goal-directed actions (and their common effects) to shared goal. Moreover, we restricted our investigation to activities for which it is the coordination among agent's actions that plays the crucial role in bringing about the collective outcome – a limitation which matches Butterfill's intention to ground joint action on the notion of "collective goals" (2011). For these goals to be recognized as such, the actions of each agent have to be mutually interpreted to understand whether "they are related to each other and to the whole event *in ways sufficient for all only the agents of these actions to be agents of the event*" (my italics). But which are these "sufficient ways" and how the agents are able to recognize them? Our modest proposal, which so far piggybacked on the structure of Butterfill's account, consists mainly in articulating an answer to this question by appealing to the teleological action understanding applied to cases involving plural activities. As Csibra & Gergely write: "an observed behavior is interpreted as an action directed to a particular end state if it is judged to be the most efficient means available to the agent for achieving this goal in the given environment" (in press; see also: Csibra, 2003). Our claim here is that, inasmuch as the principle of efficiency can be used to interpret whether an individual action represents a (rational) means to the occurrence of a goal, the same principle can be applied in interactive settings to help one agent understand how another agent's actions stand in respect to a putative shared goal. Imagine the following scene: Bob and Matt are in front of a pot into which some pasta is cooking. After ten minutes, Bob puts a pair of mittens, grabs the handles of the pot, and raises it in mid-air. Matt immediately reaches for the colander and puts it in the sink. As soon as Matt ends his action, Bob spill the boiling water and the pasta in

the colander. Assuming that both Matt and Bob know that the goal of boiling the pasta is to make it edible, and that boiled pasta can be eaten only after the water is removed, and further assuming that they know that this goal can be fulfilled by taking the pot away from the stove and spilling its content in a colander, then there is a sense in which Matt's action of putting the colander in the sink while Bob was keeping mid-air the steaming-hot pot gives Bob evidence about the fact that Matt is sharing with him the same goal of removing the water from the pasta. The role played by the principle of efficiency here is to constrain the hypothesis space about Matt's possible goal. If Bob, as we said, is able to represent the sequence of steps (i.e., means actions) needed to achieve the goal of draining pasta, then he can reason about Matt's action in terms of whether it represents an efficient means in relation to the fulfilling of Bob's goal. This would not only tell Bob that, since Matt's contribution is functional to the realization of Bob's goal, the two of them share the same goal, but also that Matt can be legitimately considered (between the many possible agents – suppose that the kitchen was crowded with people doing various things) subject of the plural activity.

There are several action parameters that Bob could have considered to interpret Matt's as directed to the shared goal – parameters that usually differ in terms of their “diagnostic value”, that is, how strongly they secure the shared-goal interpretation against other possible alternatives. While, for example, putting the colander in the sink one day before wouldn't have unambiguously informed Bob about Matt's contribution to the shared goal, the temporal contingency of Matt's action reveals a kind of responsiveness which in many cases is a robust cue of joint action (Knoblich et al., 2010). Moreover, the fact that the object selected by Matt was, among the many available at arm's reach in a kitchen, the most appropriate for a successful realization of the goal can concurrently play a role in strengthening Bob's ascription of a shared goal to Matt. Note that this kind of inferential process does not necessitate a deliberate form of reasoning, but can be carried out by implicit processes of expectation formation, which do not require attributing any kind of psychological state to the agent participating in joint action (Csibra et al., 2003). However, for goal-directed actions which involve complex structures of means actions to be carried, the parsing and sequential ordering of the necessary actions is also required. This amounts to representing what Kutz, in a more mentalistic flavor, defined executive and subsidiary intentions (2000) – the former indicating an intention whose content is an activity conceived of a whole, which “plays a characteristic role in generating, commanding or determining other intentions or mental states in order to achieve the total outcome”, and the latter being an intention whose content is the achievement of a part of the total outcome. Our account allow us also to re-describe without appeal to any psychological state the bootstrapping process envisaged by Kutz, through which non-collective elements produce a joint activity as the object of the agents' executive intention: in our scenario, this would correspond to the integration of individually selected means-selection into a plural activity.

The thorny case of the local-goal aficionados

Given the emphasis we put on what we called from the agent's viewpoint “the identification problem”, our account assumes that Butterfill's six indicators of joint action can have observable counterparts which can guide the agents' expectations about joint participation in the plural activity. However, even assuming that an agent acts in a way that is perceptually identical to that of an agent voluntarily controlling whether her actions are joint or individual (third indicator) and coordinating her action in a way that direct them jointly with the other agents' actions to the goal (fourth indicator), there is still an eventuality for which these two indicators are not sufficient to ascribe a shared goal to the agent. In fact, rather than acting in function of the shared goal, the agent's goal may have been in fact to realize of the outcome of the means action itself. The means actions, rather than being the local goal in function of the shared (global) world, is in fact for the agent the only goal represented (e.g., Bob assumed that Matt shares the goal of draining the pasta because

his action of putting the colander in the sink plays a decisive role in realizing that outcome while, in fact, Matt wanted simply to put that tool in the sink for fun). Without Butterfill's sixth indicator – i.e., the agents' disposition to adapt their actions to each other (2012) – our account would not be immune from this kind of misinterpretation. Moreover, as Kutz noted (2000) it is not immediately obvious that we can eliminate these cases from the set of cooperative activities that may count as authentic instances of joint action. "An agent, in fact, can be non-accidentally related as means to end, and the agent can intend that that relation obtain, without intending that the end be realized". What Kutz is insisting upon here includes any case of collective action in which agents are alienated from the end to which they contribute (but not because of coercion). By treating these cases as instances of joint action, his account can accommodate intentional participation by "cognitively vague" or alienated agents (2000).

We may dismiss them as fringe cases, but these rather counter-intuitive scenarios of "misattribution of shared goals" between agents represent a recurrent obstacle to overcome in the adult-infant structuring of joint interaction: "infants appear largely to be representing and pursuing their own ends in the context of this object-directed joint activity, while adults help them transform these individual pursuits into cooperative engagement" (Brownell, 2011). The goal and the working of the tasks are demonstrated by adults so that infants do not need to figure out or negotiate the goal. The only requirement for the child is to be able to coordinate her actions spatially and temporally in accordance with the partner's actions in relation to the shared goal. However, even the fortuitously occurrence of the relevant component action may, in fact, be performed by the child in a way that is independent from the state of the task as a function of each individual contribution. Young children, notes Brownell, "may be simply responding to the behavioral topology of the task in order to achieve their own goal" rather than subordinating their actions to the shared goal (2011). That is, even if infants do not have to extract the structure of the task, they may still have difficulties in representing the concatenation of individual agents' contributions to the shared goal. By showing to the children how their object-directed action are nested in a sequence of subsequent actions involving another agent (possibly through some kind of "motionese", even if the evidence about its role in facilitating imitation of action sequences in toddlers is sparse: Brand et al., 2002; Koterba & Iverson, 2009), adults can scaffold children's understanding of how (a) specific spatio-temporal patterning of actions bring about a goal, and (b) they can be distributed between agents.

Should we consider the above scenario a legitimate case of joint action relatively to our definition based on shared goals? Our answer is no. This is because the seeming coordination of the child's action, produced by a fortuitous "good fit" of its execution within the event structure that grounds it (see Pietroski's notion of grounding in Butterfill, 2011), would unravel in case of any change in the context requiring the action to be adapted to the new environmental constraints for the shared goal to be realized (cf. with Butterfill's sixth indicator). Note that here the child is not representing anything of the grounding event besides her action's outcome. This case can be compared to another which, despite featuring a shallow kind of spatio-temporal interlocking of the agents' action, still fails to meet our conditions for joint action. Suppose that Matt puts the colander in the sink right after Bob, but Matt's action is driven only by the fulfillment of an "if-then" rule involving more than one agent (e.g., if he raises the pot, I put the colander). Here there is no need to evoke the principle of efficiency to explain how the two actions (and thereof their representations) are bind together. Even assuming that Matt is not blindly performing his action, but has control over its execution and intend to "do Y if and only if Bob does X", this description still lacks any reference to shared goals.

A tale of partially shared goals

From the cases discussed so far it thus seems that the ascription of shared goal cannot be circumvented for any action to qualify as joint. But to what extent the collective goal has to be shared? Our answer, besides the unnecessary attribution of psychological states such as intentions, echoes Kutz's: "agents' intentions overlap – they share goals – when the collective end component of their participatory intentions refers to the same activity or outcome and when *there is a non-empty intersection of the sets of states of affairs satisfying those collective ends*" (Kutz, 2000; my italics). Suppose again that Matt puts the colander in the sink with perfect responsiveness to Bob's previous action. He selects the right tool and uses it at the right time; moreover, in the cases Bob's action is subjected to, for example, an unanticipated delay, Matt is able to flexibly change the timing of his action. Furthermore, Matt engages in teleological reasoning and applies the principle of efficiency to analyze whether Bob's action is the rational means to the assumed shared goal. However, Matt thinks about the joint action as something akin to the hole-fitting game children play with different shapes – he assumes that the shared goal is to put the pot in what, given the pot's shape, is the best container. Obviously, this isn't the shared goal Bob has in mind, but by observing Matt's perfectly timed action, Bob is "justified" in holding the expectation that Matt shares the collective goal of drying the pasta with him. By looking back at Kutz's quote, we can easily explain why we qualify this case as a genuine instance of shared actions: it includes a mutual ascription of shared goals which, at least within a specific segment of the action sequence, partially overlap. To understand the implications of the inclusion of this case, we will briefly compare it to the "social tool hypothesis" as described by Hamann et al. (2012). In her study, Hamann showed that 2-year-olds (but not 3-years-old) acting diadically on an apparatus cease acting together with their partner as soon as they obtain the reward (marbles that can be used to activate a jingle machine) made available by the joint pulling of two handles. Hamann – and similarly Warneken et al., 2012 – uses this evidence to show that it is not until the third birthday that children start to appreciate the normative dimensions of collaborative activities. This, in turn, suggests that in these apparently cooperative enterprises children may be pursuing the goal individually, using the partner as means to their own end – that is, a "social tool". Alternatively, Brownell (2011) provides a deflationary reading of the finding which sidesteps any normative consideration: 2-year-olds' joint action reflects reasoning about how to coordinate their behavior with another to achieve their own ends "without constructing a shared representation under which they work together to achieve a goal they hold in common". We find both readings unconvincing, for opposite reasons – Hamann's, because expectations about equal entitlement of a collective goal's outcome do not necessarily follow from the "rational pressure" of meshing cooperatively-neutral actions between agents (see: Bratman, 1993); Brownell's, because she "grants too little" to children's understanding of joint action in this particular case. In fact, by appealing to our notion of mismatched collective goals, we can still defend an interpretation of children's dyadic working on the apparatus as genuinely joint. The fact that child A ceases to pull as soon as he obtains the marbles, in fact, can presuppose the representation of a shared goal in the sense of a "non-empty intersection of the sets of state of affairs satisfying the collective ends" (Kutz, 2000). Child A may interpret child B's actions as directed to the goal of "making a subset of the marbles available to a child", whereas child B works under the assumption that A's goal is to "make all the marbles available to the children". Despite the mutually mistaken ascription of executive goals (in the sense of Kutz, 2000), the pulling actions of the two children are governed by a shared representation of the interlocking sequence of means action required to bring about an outcome, whose representation differs between the two children. As in the case of Matt and Bob playing (from Matt's viewpoint) the hole-fitting game with the colander, also in this case the two children are acting on the basis of mutual expectations about individual goal-directed actions in a way that would exhibit both Butterfill's (2012) and Kutz (2000) indicators of joint action during the execution of the task. The actions performed by each child in response

to her partner's behavior afforded a mutual interpretation of the individual actions as directed to a shared goal, because they relate to the realization of the outcome in a way that satisfies an efficiency-based action analysis of the task. The fact that the shared goals represented by the two children actually diverged in the representation of the expected outcome should not overshadow that their overlapping allowed a flexible, efficient and – most importantly – non-accidental (because based on mutual expectations about rational actions) form of coordination. Taking the comparative step, this claim holds true for all the documented cases of coordination between individuals of the same species (chimpanzees: Hirata & Fuwa, 2007; hyenas: Drea & Carter, 2009; elephants: Plotkin et al., 2012) – achieved by operating on apparatuses requiring joint, synchronous pulling to obtain food reward – which reveal an unambiguous understanding of the working of the task and on the partner's role (as showed by the fact that the animals waited for their partner to reach their position before pulling, and didn't bother trying alone when the apparatus was modified in such a way – for example, by making one of the rope ends out of the animal's reach – to make impossible for the other partner to engage in joint pulling). Saying that some animals may use the other partner as "social tool" represents a statement about the normative dimension of joint action – e.g., whether jointly acting creates collective entitlement of the outcome produced in chimpanzees: see Hamann et al., 2011 – which in no way undermines the evidence about their ability to engage in forms of joint action substantial enough to satisfy our criteria.

In conclusion, our aim was to show how it is possible for two (or more) agents to engage in relatively rich forms of joint actions without the need of prior commitments, communication, and ascription of mental states. The one and only cognitive pre-requisite we equipped our putative agents with consists in the ability to engage in teleological action understanding, and to apply its built-in efficiency-based analysis of actions to cases involving plural activities. Given the role played in our account by shared goals, we also explored to what extent goals can be shared (or not) without comprising the joint-ness of the interaction. This allowed us to finally reconsider some counter-intuitive cases of partially grounded coordination as genuine cases of joint action.

References

- Bacharach, M. (2006). *Beyond individual choice. Teams and frames in Game Theory*. N. Gold, & R. Sudgen (Eds.), Princeton: Princeton University Press.
- Brand, R. J., Baldwin, D. A., & L. A. Ashburn (2002). Evidence for motionese: modifications in mothers' infant-directed action. *Developmental Science*, 5, 72–83.
- Bratman, M. E. (1992). Shared cooperative activity. *The Philosophical Review*, 101, 327-41.
- Bratman, M. E. (1993). Shared intention. *Ethics*, 104, 97-113.
- Bratman, M. E. (2009). Modest sociality and the distinctiveness of intention. *Philosophical Studies*, 144, 149-65.
- Brownell, C. A. (2011). Early developments in joint action. *Review of Philosophy and Psychology*, 2, 193-211.
- Butterfill, S. (2011, manuscript). What is joint action? A modestly deflationary approach.
- Butterfill, S. (2012). Joint action and development. *Philosophical Quarterly*, 62, 23-47.
- Carpenter, M. (2009). Just how joint is joint action in infancy? *Topics in Cognitive Science*, 1, 380-92.

- Csibra, G. (2003). Teleological and referential understanding of action in infancy. *Philosophical Transactions: Biological Sciences*, 358, 447–458.
- Csibra, G., Biro, S., Koos, O., & Gergely, G. (2003). One-year-old infants use teleological representation of actions productively. *Cognitive Science*, 27, 111-33.
- Csibra, G. & Gergely, G. (in press). Teleological understanding of actions. In M. R. Banaji & S. A. Gelman (Eds.), *Navigating the Social World: What Infants, Children, and Other Species Can Teach Us*. Oxford University Press.
- Drea, C. M., & Carter, A. N. (2009). Cooperative problem solving in a social carnivore. *Animal Behavior*, 78, 967-977.
- Gilbert, M. P. (1990). Walking together: A paradigmatic social phenomenon. *Midwest Studies in Philosophy*, 15, 1-14.
- Gilbert, M. (1992). *On Social Facts*. Princeton, NJ: Princeton University Press.
- Gräfenhaim, M., Behne, T., Carpenter, M., & Tomasello, M. (2009). Young children's understanding of joint commitments. *Developmental Psychology*, 45, 1430-43.
- Hamann, K., Warneken, F., & Tomasello, M. (2012). Children's developing commitments to joint goals. *Child Development*, 83, 137-45.
- Hamann, K., Warneken, F., Greenberg, P., & Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees. *Nature*, 476, 328-331.
- Hirata, S., & Fuwa, K. (2007). Chimpanzees (*Pan troglodytes*) learn to act with other individuals in a cooperative task. *Primates*, 48, 13-21.
- Knoblich, G., Butterfill, S., & Sebanz, N. (2010). Psychological research on joint action: Theory and data. In B. Ross (Ed.), *Psychology of Learning and Motivation*, volume 51. Academic Press.
- Koterba, E. A., & Iverson, J. M. (2009). Investigating motionese: the effect of infant-directed action on infants' attention and object exploration. *Infant Behavior and Development*, 32, 437–44.
- Kutz, C. (2000). Acting together. *Philosophy and Phenomenological Research*, 61, 1-31.
- Liebal, K., Behne, T., Carpenter, M., & Tomasello, M. (2009). Infants use shared experience to interpret pointing gestures. *Developmental Science*, 12, 264–271.
- Moll, H. & Tomasello, M. (2007). Cooperation and human cognition: The vygotskian intelligence hypothesis. *Philosophical Transactions of the Royal Society B*, 362, 639–648.
- Pacherie, E. (2011). Framing joint action. *Review of Philosophy and Psychology*, 2, 173-192.
- Plotkin, J. M., Lair, R., Suphachoksahakun, W., & de Waal, F. B. M. (2012), Elephants know when they need a helping trunk in a cooperative task, *Proceedings of the National Academy of Sciences*, 108, 5115-5121.
- Tollefsen, D. (2005). Let's pretend! Children and joint action. *Philosophy of the Social Sciences*, 35, 75-97.
- Sugden, R. (2003). The logic of team reasoning. *Philosophical Explorations*, 6, 165–181.

Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks*, 23, 998-1003.

Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development*, 77, 640–663.

Warneken, F., Gräfenhain, M., & Tomasello, M. (2012). Collaborative partner or social tool? New evidence for young children's understanding of joint intentions in collaborative activities. *Developmental Science*, 15, 54-61.