2013

Joint attention in joint action

Anika Fiebich
Ruhr-Universitat Bochum, Germany

Shaun Gallagher
University of Memphis

Publication Details

Fiebich, A. and Gallagher, S. 2013, 'Joint attention in joint action', Philosophical Psychology, vol. 26, no. 4, pp. 571-587.

Joint attention in joint action

Abstract
In this paper, we investigate the role of intention and joint attention in joint actions. Depending on the shared intentions the agents have, we distinguish between joint path-goal actions and joint final-goal actions. We propose an instrumental account of basic joint action analogous to a concept of basic action and argue that intentional joint attention is a basic joint action. Furthermore, we discuss the functional role of intentional joint attention for successful cooperation in complex joint actions.

Keywords
attention, action, joint

Disciplines
Arts and Humanities | Law

Publication Details
Fiebich, A. and Gallagher, S. 2013, 'Joint attention in joint action', Philosophical Psychology, vol. 26, no. 4, pp. 571-587.

Joint attention in joint action[1]

Anika Fiebich and *Shaun Gallagher*

Abstract

In this paper, we investigate the role of intention and joint attention in joint actions. Depending on the shared intentions the agents have, we distinguish between *joint path-goal actions* and *joint final-goal actions*. We propose an *instrumental account of basic joint action* analogous to a concept of basic action and argue that intentional joint attention is a basic joint action. Furthermore, we discuss the functional role of intentional joint attention for successful cooperation in complex joint actions.

Keywords: joint attention, interaction, basic action, joint action, shared intention

Joint actions are frequent in our everyday life. They range from a couple going out for a walk, to children playing tag, to more complex shared and cooperative activities sometimes involving a significant number of agents and complex institutional frameworks. Bratman (1992, p. 327) mentions the example of a symphony orchestra following its conductor. Keeping things simple, we consider in the present paper joint actions that involve pairs of participating agents. Specifically we discuss the role of intention and joint attention in joint actions. Although joint attention is often discussed in the joint action debate (Knoblich and Sebanz 2008; Sebanz et al 2006; Tomasello et al 2005), a systematic account of the role of joint attention in joint actions and the conceptual relations between these concepts is still missing in the current literature. The present paper aims to provide the beginning of such an account. We propose an *instrumental account of basic joint action* analogous to a concept of basic action and argue that intentional joint attention is a basic joint action, which transitions from dyadic interaction to joint action. That joint attention is considered a basic joint action means that (1) it fulfills the minimal conditions to be a joint action, and (2) it is involved in many but not all complex joint actions. With reference to the agents' shared intention we also distinguish between two different types of complex joint actions: joint path-goal actions and joint final-goal actions.

The idea that intentional joint attention is a basic joint action (we argue) helps to clarify what can count as a joint action (i.e. how minimal one can go) and precisely when and how joint attention is part of more complex joint actions. Just as the discussion of

[1] AF acknowledges the École Normale Supérieure in Lyon, the Institut Jean Nicod (IJN) and the Barbara Wengeler Foundation in Paris for supporting research stays in France (2010); SG acknowledges support provided by research grants from CNRS as visiting professor at the École Normale Supérieure de Lyon, and the Centre de Recherche en Epistémelogie Appliquée, École Polytechnique, Paris (2010), and ongoing support from the Marie Curie project, *Towards an embodied science of intersubjectivity* (TESIS). The final draft of this paper was completed with support from the Humboldt Foundation Anneliese Maier Research Award.

individual basic action in action theory attempts to say what kinds of things can count as actions, so in our analysis we attempt to say what kinds of things can count as joint actions. Furthermore, our account of joint action does not focus on the cognitive stance of a single individual (the kind of 'methodological individualism' that one finds in Searle, for example) but rather on what joint action implies about the interaction of two (or more) agents engaged in a joint action and involved in action coordination. To motivate this perspective we start with De Jaegher et al.'s (2010) notion of interaction.

Interaction, shared intentions and joint action

Following De Jaegher et al. (2010), we take interaction to be a mutually engaged co-regulated coupling between at least two autonomous agents where the co-regulation and the coupling mutually affect each other, and constitute a self-sustaining organization in the domain of relational dynamics. If the autonomy of one agent is dissolved, because one agent is the sole regulator of the coupling and the other is just co-present, this is no longer interaction in the strict sense. An agent's relation with an object or tool is also not an interaction since here too there is no mutuality. As long as the conditions of this definition are fulfilled however, it can apply to cross-species interactions or even interactions with robots (that are autonomous in the sense intended). In the present paper we limit considerations to dyadic human-human interactions ('social interactions'). In general, interactions include social interactions, and social interactions include joint actions.

Joint action is a complex form of social interaction. There are different views in the current literature about exactly what constitutes joint action. Sebanz, Bekkering, and Knoblich (2006) define joint action as "any form of social interaction whereby two or more individuals coordinate their actions in space and time to bring about a change in the environment" (p. 70). They identify joint attention, action observation, task-sharing and action coordination as cognitive mechanisms through which a successful joint action can be achieved but do not suggest that shared intentions are involved in joint actions. On our view, in contrast, the coordinated behavior patterns are (more or less) specified depending on the shared intention the agents have. Joint actions always involve shared intentions (Carpenter 2009), and in part this is what distinguishes human group activities from those of other animals (Tomasello et al 2005).

To clarify the notion of shared intention, consider first that *common knowledge* of aiming for the same goal is a crucial requirement for joint actions, as indicated by Sebanz et al (2006) and Tomasello et al (2005). Not only do we have the same goal, we are aware that we have the same goal, which makes it a *common goal*. But the fact that two individuals each have the personal intention to aim for a particular goal and the fact that they mutually know that they are both aiming for the same goal are not sufficient to capture the kind of intention involved in a joint action. Additionally, the personal intentions of the two agents need to involve a specific 'we'-activity to make it a shared intention had by both agents. Searle's (1990) concept of 'we-intentions' can help to specify the nature of shared intentions in joint actions. For Searle, joint action is collective intentional behavior involving we-intentions. Collective intentional behavior is not the same as the summation of individual intentional behaviors; thus, we-intentions cannot be analyzed into sets of I-intentions, even if these I-intentions are supplemented with beliefs about the intentions of others. If two (or more) agents aim for a common

goal in a way that their we-intentions are non-reducible to a set of I-intentions, we can say they have a 'shared intention'.

> Imagine that a group of people are sitting on the grass in various places in a park. Imagine that it suddenly starts to rain and they all get up and run to a common, centrally located shelter. Each person has the intention expressed by the sentence 'I am running to the shelter'. But for each person, we may suppose that his or her intention is entirely independent of the intentions and behavior of others. In this case there is no collective behavior; there is just a sequence of individual acts that happens to converge on a common goal. Now imagine a case where a group of people in a park converge on a common point as a piece of collective behavior. Imagine that they are part of an outdoor ballet where the choreography calls for the entire *corps de ballet* to converge on a common point. We can imagine that the external bodily movements are indistinguishable in the two cases; the people running to the shelter make the same types of bodily movements as the ballet dancers. Externally observed, the two cases are indistinguishable, but they are clearly internally different."
>
> (Searle, 1990, pp. 402-403)

In Searle's first case, each person has the same goal (i.e. running under the shelter) but each one could express his or her intention to achieve that goal without reference to the others, even if each has common knowledge of the intentions of the other. They have a *common* goal, which is more than having the same goal since mutual knowledge of aiming for that goal is involved. In Searle's second case, the dancers have what we call a 'shared intention'. On our view, having a shared intention implies not just having a common goal but also having a common goal that involves we-intentions (see table 1). The notion of we-intention implies the notion of cooperation and this involves interaction as defined above. Simply having mutual knowledge of having the same goal is not sufficient for a we-intention; it also requires cooperation among the agents or intention to cooperate among the agents.

Crucially, however, for Searle, we-intentions are not shared intentions. He neither uses the term 'shared intention' nor makes use of the concept of we-intention to account for such a phenomenon. According to Searle, we-intentions are attitudes of single individuals. They are special intending attitudes of single individuals with a special mental content that involves a 'we'-activity.

In contrast, Bratman provides an account of shared intention in which he highlights the idea that a shared intention consists primarily of the interrelations of the agent's attitudes. Following Bratman (1993), we take a shared intention "not as an attitude in any mind. It is not an attitude in the mind or minds of either or both participants. Rather, it is a state of affairs that consists primarily in attitudes (none of which are themselves the shared intention) of the participant and [intersubjective] interrelations between attitudes" (pp. 107/8). This approach needs to be distinguished from the atomism of Searle's account of 'collective intentions', as well as from non-individualistic accounts such as Roth's (2004) account of shared agency and contralateral commitments, or Gilbert's (2009) 'plural subject account' according to which agents form a joint commitment to act as a body as would a single individual. Bratman (2009) suggests that "a Searlean we-intention is, then, a candidate for the intentions of individual participants that together help constitute a shared intention, though Searle himself does not say how the we-intentions of different participants need to be inter-related for there to be a shared intention" (p.41).

In addition to having shared intentions, Bratman (1992) adds three other features to joint action. First, he suggests that mutual responsiveness to the intentions and actions of the other is necessary. This, however, is a feature of any kind of interaction and not specific to joint action. He also adds, second, commitment to the joint activity, and commitment to mutual support of the efforts of the other to play her role in the joint activity. Although we acknowledge a point made by Gilbert (2009), that personal readiness needs to be expressed by the agents for a joint commitment to the shared intention (e.g. by verbal or non-verbal gestures) in order to become common knowledge among the agents, we will argue that this does not necessarily involve reciprocal obligations of the agents to pursue the action or to act in a certain way. Finally, Bratman adds the idea that in joint action we need to mesh sub-goals.

Even though we agree in large parts with Bratman, we put less weight on the need to mesh sub-goals. As we will argue below, the meshing of sub-goals may, but does not need to play a role in joint actions, and in some cases there are no sub-goals. Accordingly, we bracket this requirement in what we suggest are the conditions for joint action. All three of the following conditions are necessary, and together they are sufficient for joint action:

1. having a shared intention (i.e. aiming for a common goal, involving we-intentions).

In general, having a common goal presupposes

2. having common knowledge (shared awareness) of aiming for the same goal [and sometimes being obligated to pursue and support the intention].

If agents have merely a common goal, they do not need to cooperate in order to achieve that goal (see above). However, having a shared intention means

3. participating in cooperative behavior patterns (determined to varying degrees by rules or regularities) in order to achieve the goal [which may or may not include meshing sub-goals].[2]

Whether and to what extent obligations are involved, and whether and to what extent meshing sub-goals is required, depend on the shared intention of the particular joint action. The shared intention involved in a joint action determines how agents need to coordinate their behavior in order to succeed in their cooperation. As we will show below, joint action is a complex form of social interaction in which agents do not just coordinate their behavior, but where coordinated behavior patterns are (more or less) specified in advance by rules and regularities dependent on the shared intention the agents have. 'Cooperation' involves coordinated behavior patterns, based upon those rules and regularities, to achieve a common goal.

Joint actions can be *joint final-goal actions*, where agents coordinate their behavior in order to achieve an end-product or end-state, i.e. a final goal, which can be independent from the particular coordinated behavior pattern used to achieve that final goal. That is, the goal, e.g. robbing a bank, may be achievable in any number of ways. However, agents may also coordinate their behavior as an end in itself; call this a *joint*

[2] Crucially, to fulfill condition (3), the agents do not need to interact with each other cooperatively *at the same place*; as we will see below, a set of instructions by a coordinator may subserve the cooperation of agents at different places.

path-goal action. In such cases, e.g. dancing together, the shared intention entails only the activity itself, and thus, the coordinated behavior patterns to achieve the common goal involving we-intentions are more (as in tango) or less (as in free-form disco dancing) specified in advance. The distinction between joint final-goal and joint path-goal actions should be viewed not as a strict difference in kind, but as two ends of a continuum since many actions share characteristics of both. As we'll see, in this regard joint actions are sometimes ambiguous and can be described in complex ways.

The more rule-governed the shared intention (i.e. the more the coordinated behavior patterns are specified in advance), the more specific are the commitments involved and the higher are the action expectations and obligations. It also seems right to suggest that less cognitive effort to mesh or monitor sub-goals would go into such rule-governed contexts. The sub-goals are already pre-ordered by the rules or regularities. For example, if we decide to play chess together, we commit to observe the conventional rules entailed in chess and our action expectations in playing chess are based on these commitments; thus, you would protest if I used the rook like the queen. Also, once we accept all the rules entailed in chess, no sub-goals are left that might divert or stop our joint activity since our action opportunities within the game are prescribed in advance[3]. Additionally, we are not only obliged to observe the rules of chess when we decide to play chess together but we are also committed to finishing the play (though the commitment to finish the game is less formal than following the rules of the games). Thus, if your brother, who is an excellent chess player and with whom playing chess might be much more exciting for you than playing with me, arrives, I nonetheless expect you to finish the game with me.

At least three kinds of constraints characterize the rules or regularities that determine the coordinated behavior patterns of two agents engaged in joint action in order to achieve a common goal involving we-intentions: (1) *natural constraints* (such as regularities that are biologically specified. Imagine if we decide to go for a walk, the movements of the shared activity, walking, are determined by certain biological constraints); (2) *rational constraints*[4] (i.e. aspects that are rationally specified such as mathematical rules that need to be observed in e.g. running an experiment), and (3) *social/conventional constraints* (i.e. aspects that are specified by conventions like the rules of chess but also social-cultural norms such as the agreed upon goal of finishing the game).

In many joint path-goal actions cooperative behavior is highly specified in advance by rules and regularities. Even in cases such as free-form disco dancing or improvisation, behavior is constrained by regularities and rules that define the action to be what it is, e.g., moving in relation to the music or following a general pattern within a specified realm of meaning. In contrast, many joint final-goal actions allow for different possibilities of coordination patterns useful for achieving the final-goal. Accordingly, in joint final-goal actions, more effort might be required to monitor and mesh sub-goals.

[3] There is no space left for sub-goals concerning the joint activity itself but there may be diverging sub-goals concerning the situational features (location, date, etc.) in which the joint activity takes place.

[4] A regular application of logical/mathematical rules for the same purpose may become conventional by and by.

Imagine that we decide to cook together and that we take this to be a joint final-goal action where the final goal is to have dinner as the end product of our cooking; we enjoy eating but not cooking. In this joint action there is still a lot of space left within this shared goal for diverging sub-goals; e.g., you might decide on a rare steak whereas I prefer a vegan dish. In other cases we might have different sub-goals that do not contradict each other, e.g., you might have the sub-goal to eat Spaghetti Bolognese with onions whereas I have the sub-goal that we eat at my home. In some cases, however, joint final-goal actions might also involve specifying in advance the coordinated behavior patterns to achieve the goal, which is the case if we decide to cook a specific recipe.

The more rule-governed our joint action, the more are we obliged to follow these rules and the more are we in a position to demand from one another that we follow these rules. In general, joint actions involve the basic joint commitment to pursue the shared intention until it's fulfilled. In joint final-goal action, fulfillment is achieved when we accomplish the final state or end-product for which we aimed. In joint path-goal actions, fulfillment is achieved simply by engaging in the joint action. In many joint path-goal actions fulfillment is not specified in terms of completing or bringing the action to a specific end. If so, no obligations to continue the joint action are involved in the shared intention. Imagine, e.g. we happen to meet each other in a discotheque. You pull me onto the dance floor and start to dance, and I respond to your dancing movements. Our joint intention is spontaneously formed in the action, and our commitment to the action is relatively weak. Either one of us can stop dancing to take a drink, or to start dancing with someone else whenever we want, and neither of us is in a position to expect or demand the other to continue dancing to the end of the song.

Joint attention

What, if any, role does joint attention play in joint actions? It is important to note that the terms 'shared attention' and 'joint attention' are defined in various ways in the current literature, and are sometimes used interchangeably (see Triesch et al 2006). Thus, we first need to clarify how we use these terms in dyadic and triadic interactions and joint actions.

Simple and shared attention. From birth on, infants are attentive to external entities (call this 'simple attention') and are engaged in *dyadic self-other interactions* which involve dyadic attention where subjects are mutually attending to each other. Later, when the infant begins to follow the gaze of the other person, it may occasion a new form of attention (call this 'shared attention'): the infant is aware of the adult being attentive towards the object *and* of herself being attentive towards the object. Baldwin (1995) equates joint attention with the simple result of gaze following: "the simultaneous engagement of two or more individuals in mental focus on one and the same external thing" (Baldwin 1995, p. 132). In contrast, we suggest that this co-orientation is not yet joint attention, and is only *shared* attention when one of the individuals is aware of himself and the other as being attentive towards the same external entity. Shared attention involves awareness, on the part of one subject, that both subjects are attending – I may knowingly attend to the same object as you attend to, but without you knowing it. However, shared attention, on this definition, is not yet a triadic form of interaction; it is rather a dual attending – a combination of simple attendings: a simple attending to an external entity, and a simple attending to the person being attentive towards the same

entity.

Joint attention. Other theorists define joint attention by adding a further characteristic to shared attention, i.e., the *mutual* knowledge of *both* individuals as being attentive towards the same external entity, i.e. an object, other person, or event (e.g. Tomasello et al. 2005) accomplished via communicative cues (Carpenter and Liebal, in press). Call this 'joint attention' in a *triadic self-other-entity interaction*. From 9-10 months onward, in the phenomenon of social referencing, for example, when infants start to refer to the other's emotional expression to know whether to approach novel objects (Klinnert et. al. 1986, Moses et. al. 2001), the infant is engaged not just in shared attention but in *joint* attention, i.e., an interaction that includes the awareness that she and the other are both perceptually attentive towards the same external entity *and* that the other is aware of this shared attention as well.[5] Striano and Rochat (2000) showed that whether 10-month olds (but not 7 month-olds) monitor and refer to others in ambiguous situations depends on the others' attention towards them (i.e., the infants). That communicative cues in social learning situations involving joint attention evoke different learning effects than if those cues were absent is supported by empirical evidence from developmental psychology (Csibra 2010).

Usually, joint attention means that individuals are visually mutually attentive towards an external entity. However, joint attention can also include other sensory modalities, e.g. our auditory modality when we enjoy listening to a concert together. Joint attention can be conducted in any perceptual modality, and in some cases is verified in communication (e.g., "Do you taste the sweetness of this wine?" – as we both taste the wine).

Intentional joint attention. In some of these examples, joint attention may involve a shared intention. If so, agents are not just mutually aware of being attentive towards the same entity. Rather, they also *intend* to be mutually attentive towards the same entity (where the shared intention may just be to maintain joint attention). On this definition, intentional joint attention (iJA) fulfills all three conditions of a joint action. In iJA, there is a shared intention (condition 1), which, at a minimum, may entail the immediate common goal (involving we-intentions) of maintaining joint attention (e.g., when we are both surprised by some object or event and desire to jointly remain attentive, perhaps for the sake of mutual enjoyment, or continuing our conversation). Even in the minimal case there is common knowledge (and likely desire) to maintain joint attention as the goal (condition 2), and to coordinate our behavior patterns to achieve this goal (condition 3). iJA may be involved in complex joint actions that take place in real interactive settings. Often, we need to aim for or maintain joint attention in order to cooperate. The coordinated behavior of our cooperation is defined by the shared intention we have (see previous section). Crucially, in joint actions we may require iJA either towards an external entity (e.g. if our shared intention involves lifting the box in front of us together) or to the event of our interaction itself (as in dancing the tango).

[5] Crucially, we use the term 'joint attention' to mean 'joint attention in perception'. There might be another way to use the term to include 'joint attention in imagination' which goes beyond perception and involves the conceptual in the sense that two agents are jointly attentive towards a concept or an idea (our conversation about justice, for example, requires that we mutually attend to this concept and the conversation itself confirms that we do). But this is not the focus of the present paper.

Accordingly, joint actions may occur not only in *triadic relations* (agent-agent-object) but also in situations of *dyadic interactions*.

Intentional joint attention as a basic joint action

In joint attention each of the jointly attending perceivers is aware of herself and the other being attentive together towards an entity or event. When joint attention is involved, both agents typically interact in a triadic relation (agent-agent-object/event). Joint attention itself may (but does not necessarily) elicit a shared intention-in-action, perhaps because of the environmental circumstances or because of the intentional behavior of one of the agents. If it does, it becomes a joint action. In the minimal case, the shared intention to maintain joint attention is an end in itself, as in proto-declarative joint attention. As Aristotle (*Nichomachean Ethics*, Book I) suggested, some actions are ends in themselves and are done for their own sake. Proto-declarative joint attention is often characterized precisely as something children engage in for its own sake (see Nichols, et al. 2005). Engagement in joint attention as an end in itself, i.e. as a joint path-goal action, may be performed for the sake of mutual enjoyment. In our everyday life, however, we usually engage in iJA for reasons other than mere enjoyment. As we suggest below, iJA may fulfill the functional role of keeping our joint action on track and/or avoiding failure in regard to our shared intention.

Given the definition of joint action outlined above, we argue that iJA is a *basic joint action*, on analogy with Chisholm's (1969) instrumental definition of *basic actions*. First, we briefly review the literature on the concept of *basic action*, and then provide an *instrumental account of basic joint action*.

According to Chisholm (1969), an action is basic when an agent performs it without performing it *by* performing some other action. In the case of a normal, complex, intentional action, if my intention is to get a drink, this is made up of more basic actions such as walking across the room, reaching and grasping the drink, etc. Not only are these basic actions intentional movements on my part, but also without these basic actions I would never be able to get my drink. Danto (1965), in contrast, appeals to the concept of causal relation to define the criterion of basicness. He emphasizes that "when M performs a basic action, he does nothing first that causes it to happen" (p. 142). Both of these views are controversial, as Hornsby (1980) makes clear. The first involves controversies about how to individuate actions (getting a drink may just be walking, reaching, and grasping such that these are not separate actions). In regard to Danto's definition, there is general consensus that a person's doing one thing does not *cause* her doing another (Hornsby 1980, p. 67). Moreover, various examples of basic actions provided by Danto and Chisholm suggest that simple (unmediated) bodily movements are paradigm cases of basic actions. However, this is not a consensus thesis in the literature and is denied by some theorists (e.g. Baier 1971; Hornsby 1980).

Baier (1971, 168ff) further highlights that actions might be considered basic in various ways; (a) *causally* (as proposed by Danto), (b) *instrumentally* (as proposed by Chisholm), (c) *conventionally* -- "in the sense of being less dependent on rules or context,", (d) *compositionally*, (e) *logically*, (f) *ontogenetically*, (g) in regard to ease of performance or (h) *ease of performance in isolation from the other actions*. Such criteria, moreover, may be incompatible with each other; e.g. whereas an action may be basic in a genetic sense, it need not be the easiest to perform.

The value of the notion of a basic action for understanding more complex, non-basic actions depends on the sense in which an action is considered as basic. For example, if we consider an action as basic in a compositional sense, a basic action constitutes (together with other basic actions) the component(s) of a complex joint action. To drive a nail into the wall, I need to hold the nail with one hand and to hammer on it using the other hand; 'holding the nail' and 'hammering on it' are both basic actions in a compositional sense that together constitute the complex action 'driving a nail into the wall'. If we consider an action as basic in an instrumental sense, in contrast, the basic action provides (one of) the means to performing a complex action. To ventilate the room, for example, I need to open the window, which requires me to grasp the window catch and turn it around. I perform that rotating movement in an unmediated way without performing some other action; thus, it is a basic action in an instrumental sense.

Relevant to the issue of individuating actions, one can ask whether basic actions (however defined) are really actions. Sneddon (2006) considers basic actions not to be real actions, but to be theoretical entities the existence of which requires demonstration (2006, p. 101). The usual strategy in regard to establishing the reality of basic actions is (1) to point to the phenomenology of action – i.e., that some bodily actions do seem to be unmediated and directly willed – or (2) to offer an argument involving avoidance of infinite regress – i.e., that not all actions can be mediated (see. e.g., Danto, 1979, p. 46). Sneddon, however, argues that basic actions are mere abstractions from normal complex actions instead of real actions themselves. On this view, the phenomenology, rather than supporting the existence of basic actions, only leads to reflective abstractions about action. Furthermore, the avoidance of infinite regress means, Sneddon suggests, that at bottom, instead of basic actions, we find processes that do not count as genuine actions; bodily movements should not be considered actions *per se*. Here one might accept that basic actions are only actions in a derivative sense since their intentional status, and important aspects of their motor control, derive from the complex intentional actions that they serve (see Baier's example of tying shoelaces in order to make one's hands move in the right way; also Ripley 1974).

Not all controversies about the concept of basic action, however, are relevant to the concept of a basic joint action. For example, Sneddon's suggestion that basic actions are not real does not apply to the concept of a basic *joint* action since there is no claim that we perform a basic joint action without performing some other real non-joint actions, and, unlike the concept of basic action, there is a clear motivational history that can be defined for a basic joint action such as intentional joint attention. Whereas debates about basic action may concern whether certain bodily events can be individuated as actions on their own, or not, and whether these events are necessary for and logically prior to more complex action, basic joint actions are *joint* actions as defined (above) by clear conditions involving (at least) two interacting agents and their *shared* intention. Indeed, one can analyze joint action into more subtle components of interaction (those that involve shared attention or the various bodily movements that constitute the required coordination). But, as we have shown, such components cannot be considered more *basic joint* actions, since they are not joint actions at all, even if they are interactional. What turns out to be more basic turns out *not* to be more basic *joint* actions, but non-joint actions (e.g., bodily movements of coordination) and interactions without shared intentions. Thus, we can offer an *instrumental account of basic joint action*:

> A joint action is a basic joint action if two (or more) agents perform the joint action without performing it by performing some other joint action.

Crucially, a basic joint action is prior to the distinction between joint path-goal actions and joint final-goal actions and can be a basic part of either kind.

On this account, we argue that intentional joint attention is a basic joint action. Intentional joint attention is a real interactional process (not a theoretical entity or mere abstraction from some more complex action) in which we engage with others. Since iJA, insofar as it involves (1) a common goal involving we-intentions, (2) common knowledge of intending that goal, and (3) participation in cooperative behavior patterns in order to achieve this goal, is a joint action, and, on analysis, its components are not themselves other joint actions (but rather more subtle components of interaction – those that involve shared attention or the various individual actions that constitute the required coordination), then it is a basic joint action, and in many (and perhaps most) cases of more complex joint action, it operates as such.

The value of the notion of basic joint action in an instrumental sense for understanding a complex non-basic joint action is analogous to the value of the notion of a basic action in an instrumental sense for understanding a complex non-basic action (see above). A basic joint action such as iJA provides (one of) the means for performing a complex joint action. In the following section, we illustrate how iJA as a basic joint action provides a means to perform a complex joint action. In some cases, iJA is necessarily involved as a means to fulfill the shared intention of the complex joint action (e.g., if we intend to help each other lift a box onto a truck). In other cases, iJA may (but does not need to) be included in complex joint actions as a basic joint action (e.g., when we cook a recipe).

Whereas joint attention may be involved in triadic interactions, *intentional* joint attention (i.e. joint attention for a mutually acknowledged purpose) is, as we will show, involved in many joint actions as a basic joint action in an instrumental sense. There are also cases of joint action, however, in which no joint attention is involved at all (see the example of terrorist action below). We have defended an *instrumental* and not a *compositional* account of basic joint action, and this is crucial, since iJA as a basic joint action may but does not need to be involved in complex joint actions such as, for example, playing chess (see below). Furthermore, iJA can stand on its own as something done for its own sake, as in proto-declarative joint attention. Insofar as iJA does play a role in complex joint actions that take place in real interaction settings, iJA serves purposes other than itself, e.g., it may help to prevent failure of the larger action or to improve performance; accordingly, it derives its significance from the larger, more complex shared intentions and joint actions that it serves.

Complicating the space of joint actions

We made a distinction between (1) *joint path-goal action*, and (2) *joint final-goal action*. As we have seen above, this distinction relates to the question of whether the coordinated behavior patterns of the joint action in question need to be specified in advance. Furthermore, this distinction has motivational implications; whereas we are

motivated to perform joint final-goal actions because we aim to achieve a certain end-product or end-state, we are motivated to perform a joint path-goal action as an end in itself, i.e. for the sake of mutual enjoyment.

As we indicated above, the distinction between joint final-goal and joint path-goal actions should be viewed not as a strict difference in kind, but as two ends of a continuum. In this regard joint actions are sometimes ambiguous and can be described in complex ways. At either end of the continuum we *both* may aim to perform a joint action (i) for its own sake, as a joint path-goal action (e.g., for the enjoyment of dancing) or (ii) in order to achieve a certain end-product or end-state, as a joint final-goal action (e.g., to win a dance contest). Furthermore, in joint final-goal actions, both of us may also enjoy the activity itself, in which case the joint final-goal action may involve sub-path-goals. It is also possible that only one of us enjoys dancing for its own sake while the other is only in it for the win. Whether our sub-goals are in agreement or not, we may both agree on the final goal, in which case the joint action remains a joint final-goal action defined as such by the same end-state or end-product that both agents aim to achieve.

Joint final-goal actions and joint path-goal actions are distinguishable according to such motivational aspects even as they may be indistinguishable to an external observer; our dancing the tango may look the same regardless whether we are dancing it for the sake of mutual enjoyment (as a joint path-goal action) or to improve our performances in order to win at the next dance contest (a joint final-goal action). In many cases, however, motivations may be mixed. Along the continuum, various possible combinations test the limits of the concept of shared intention and the 'jointiness' of the action. I may aim to dance the tango with you to win the contest, whereas you dance it only because you enjoy it (i.e., as a path-goal). Or winning the contest may be only your sub-goal, and your real intention is to impress the person you love, who is in the audience, whereas I don't care if we win the contest as long as we prevent my arch enemy from winning. Many of the joint actions we perform in everyday life may be situated somewhere in the middle area of this continuum.

In addition, iJA may or may not be involved in complex joint actions, whether joint path-goal actions or joint final-goal actions. A necessary condition for iJA being involved in a complex joint action is that the joint action in question takes place in a real interactive setting. As we will see, joint actions do not need to be performed in real interactive settings, and if so, it may be that no joint attention is involved at all. To determine whether iJA is involved or not in a particular joint action, further situational aspects need to be taken into account. We can begin by asking whether, in any particular circumstance, iJA could perform its main functional role[6] in joint actions, i.e., *to reduce the risk that the shared intention fails*. As we have shown above, our cooperation in any joint action involves coordinated behavior patterns to achieve a shared goal – patterns which are based upon a set of natural, logical and/or conventional constraints of rules and regularities that are more or less specified in advance. In order to succeed in our cooperation, we aim to avoid the possibility that our shared intention will fail. Whether iJA is involved in a particular joint action or not in order to serve that functional role, depends upon the situational features in which the joint action takes place. Furthermore,

[6] We do not claim that this is the only functional role of iJA in joint actions. Another functional role of iJA in a particular joint action might be to improve our performance in a real interactive setting. In order to improve our performance in tango, e.g., we may need to be jointly attentive towards our actions.

these situational features might determine to what degree iJA will depend upon the rules and regularities of the joint action in question.

Let's start with an example of an action that involves a shared intention. On my way to the circus I happen to meet an acquaintance on the street. She is heading to the clown school and so is walking in the same direction. While she may or may not say "I'll join you," there is an implicit intention that we will walk together just in the fact that she does join me. This sets up a joint action of walking together. We adjust our gait in order to stay together as we walk. However, we can adjust our gait without jointly attending to our walking. If an obstacle occurs on the road, in contrast, we may intentionally enter into joint attention in regard to this obstacle depending on whether in the given situation iJA is required to avoid the obstacle in a coordinated fashion. We cooperate and coordinate our behavior in a way that is determined by the natural aspects of our bodies and the environment, with or without iJA being required.

In more complex forms of joint actions, stronger conventional commitments may be involved, e.g., in playing a conventional game like chess. However, our commitment to observe the rules of chess does not necessitate iJA. In playing chess, we may intentionally jointly attend to each move, and joint attention may be involved in almost the whole interaction. However, joint attention towards our moves is not required to fulfill our shared intention. We could play a game of chess without simultaneously being in the same room where the chessboard is located. For example, at noon I make a move on the chessboard while you are away eating lunch. I then go to lunch and in my absence you return to the chess board, recognize my move, make your own, and then leave before I return, etc. etc. It would be difficult to claim that we engaged in joint attention at all, even though we could be said to be engaged in the joint action of playing chess. In this case, what keeps the game on track, and what mitigates the risk of failing to complete the action, are the established and agreed upon rules of chess playing.

In other joint actions, successful cooperation cannot be achieved without joint attention. If so, iJA is necessary to fulfill the shared intention. For example, if we intend to help each other lift a box onto a truck. In other cases, a certain use of language may substitute for iJA. Imagine, for example, we set out to cook a certain dish together and follow a recipe. Joint attention is required to keep cooperative behavior patterns on track to achieve the goal. We need to be jointly attentive towards those actions which determine the success or failure of our joint action, for example, adding ingredients to the pot. If you didn't see me salt the soup, you would run risk of salting the soup a second time. In this example, a glance into the pot (in contrast to a glance at the chessboard) is not sufficient to determine the other's action in order to coordinate our behavior. Hence, in following the recipe we intend to be jointly attentive towards our actions, and that, in part, determines the success or failure of our joint action. If our iJA is interrupted for some reason, we are committed to inform each other by verbal or nonverbal gestures about those actions we carry out that are essential for our joint action (which implies the commitment to pay attention to each other's attention as well).

As a substitute for iJA, however, we could employ a complete set of written instructions that specify who should salt the soup, and when, etc. In that case, when I am attending to my task of salting the soup at noon, you may be reading your instructions about stirring the soup at 12:01, and we may proceed through the entire project without jointly attending to the other person's actions. As the following example makes clear, language may take over the functional role of intentional joint attention and

allow for successful cooperation in joint actions not just in real interactive settings but also in joint actions that do not involve any interaction among the agents at all.

Imagine two terrorists having the shared intention of blowing up Big Ben, each one knowing that they share the same goal, and each one knowing that the goal will be reached only in cooperative actions. Neither terrorist, however, knows precisely what the other one will be doing. They wait at different places for instructions from their controller, and their behaviors are coordinated by his instructions. Here, no joint attention is involved; since one does not interact *directly* with the other and each is located in a different place, each is attentive only towards his own activity and no visual joint attention towards the other's activity is possible. In this case, to carry out the joint action, a set of instructions from the controller is substituted for the iJA that in other cases (as in real interactive settings) might be required to succeed in achieving the common goal involving we-intentions. Nonetheless, the three defining conditions of joint action are fulfilled.

In general, in joint actions there are matters of degree and a variety of ambiguities concerning the formation of shared intentions, the functional role of iJA, and the type of joint action involved. Joint actions can be planned in advance, and if so, they often involve certain spatiotemporal coordinates – when and where the action is to be performed. But joint actions can also emerge spontaneously, from a spontaneous joint attention due to the environmental circumstances of the social interaction or the intentional behavior of one of the agents.

In some situations if formation of the shared intention and the common knowledge of having that intention are not prior to the activity itself, it would be paradoxical if the agents were able to cooperate, i.e. to coordinate their behavior to achieve their common goal *together*. We are not able, for example, to play Beethoven's Ninth Symphony together in an orchestra without having the prior intention to do so (and the musical capabilities, of course). In other situations, however, prior shared intentions are not required. For example, if someone simply grabs another person and starts to dance, the shared intention – to dance – only emerges in the already ongoing action. The other person says (or thinks), after the fact, "Ok, let's dance" Or "Hey, we're dancing!" Only at that point does the shared intention emerge. Cooperative behavior patterns themselves however involve shared intentions-in-action. To carry out those shared intentions-in-action, joint attention may be required, or some set of instructions may substitute for it. Either by practice involving joint attention, or by following such instructions, joint actions may become habitual, or more formally established in or as institutions.

Summary and conclusion

In this paper, we proposed a concept of joint action according to which the nature of the shared intention involved determines how agents coordinate their behavior in order to succeed in their cooperation. In this regard, we distinguished between (i) *joint path-goal actions*, and (ii) *joint final-goal actions*, not as a difference in kind but as two ends of a continuum. We discussed different constraints involved in rules and regularities, which determine cooperative behavior patterns. Such rules and regularities lead to higher normative expectations but to a lower demand for meshing sub-goals. Furthermore, we proposed an *instrumental account of basic joint action* analogous to a concept of basic

action, and argued that intentional joint attention is a basic joint action in an instrumental sense that may – but does not need to be – involved in complex joint actions. We discussed the functional role of intentional joint attention in joint actions, namely, to reduce the risk of the shared intention failing. Using a number of examples, we illustrated that the situation in which a particular joint action is performed needs to be taken into account in order to determine whether or not intentional joint attention is involved – or whether language, some set of instructions, or certain institutional arrangements can take over its functional role and substitute for intentional joint attention.

References

Baier, A. (1971). The search for basic actions. *American Philosophical Quarterly,* 8(2), 161-170.

Baldwin, D. A. (1995). Understanding the link between joint attention and language. In C. Moore & P. J. Dunham (Eds.), *Joint Attention: Its Origin and Role in Development* (pp. 131-159). Hillsdale, N. J.: Lawrence Erlbaum.

Bratman, M.E. (1992). Shared Cooperative Activity. *The Philosophical Review*, 101(2), 327-341.

Bratman, M.E. (1993). Shared Intention. *Ethics*, 104(1), 97-113.

Bratman, M.E. (2009). Shared Agency. In C. Mantzavinos (Ed.), *Philosophy of the Social Sciences: Philosophical Theory and Scientific Practice* (pp. 41-59), Cambridge: Cambridge University Press.

Carpenter, M. (2009). Just how joint is joint action in infancy? *Topics in Cognitive Science*, 1, 380-392.

Carpenter, M., & Liebal, K. (in press). Joint attention, communication, and knowing together in infancy. In A. Seemann (Ed.), *Joint Attention: New Developments in Philosophy, Psychology, and Neuroscience.* MIT Press.

Carpenter, M. (2009). Just how joint is joint action in infancy? *Topics in Cognitive Science*, 1, 380-392.

Chisholm, R. (1969). Some puzzles about agency. In K. Lambert (Ed.), *The Logical Way of Doing Things* (pp. 199-217), New Haven: Yale University Press.

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language*, 25, 141-168.

Danto, A. (1965). Basic Actions. *American Philosophical Quarterly,* 2(2), 141-148.

Danto, A. (1973). *Analytic Philosophy of Action*. Cambridge: Cambridge University Press.

Danto, A. (1979). Basic actions and basic concepts. *Review of Metaphysics* 32 (3): 471-85.

Davidson, D. (2001). *Essays on Actions and Events*. Oxford: Oxford University Press.

De Jeagher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences*, 14(10), 441-447.

Gilbert, M. (2009). Shared intention and personal intention. *Philosophical Studies*, 144, 167-187.

Hornsby, J. (1980). *Actions*. London: Routledge and Kegan Paul.

Klinnert, M.D., Emde, R.N. Butterfield, P., & Campos, J. J. (1986). Social referencing: The infant's use of emotional signals from a friendly adult with mother present. *Developmental Psychology,* 22 (4), 427-432.

Knoblich, G. & Sebanz, N. (2008). Evolving intentions for social interaction: from entrainment to joint action. *Phil. Trans. R. Soc. B*, 363(1499), 2021-2031.

Moses, L. J., Baldwin, D. A., Rosicky, J. G., & Tidball, G. (2001). Evidence for referential Understanding in the emotions domain at twelve and eighteen months. *Child Development*, 72, 718–735.

Nichols, K. E., Martin J. N., & Fox, N. A. (2005). Individual differences in the development of social communication: Joint attention and temperament. *Cogniţie, Creier, Comportament /Cognition, Brain, Behavior*, 9(3), 317-328.

Ripley, C. (1974). A theory of volition. *American Philosophical Quarterly*, 11 (2), 141-47.

Roth, A.S. (2004). Shared Agency and Contralateral Commitments, *The Philosophical Review*, 113(3), 359-410.

Searle, J. (1990). Collective Intentions and Actions. In Cohen, P., Morgan, J. and Pollack, M. (Eds.), *Intentions in Communication* (pp. 401-415). Cambridge, MA: MIT Press.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, 10(2), 70-76.

Sneddon, A. (2006). *Action and Responsibility*. Dordrecht: Springer.

Striano, T. & Rochat, P. (2000). Emergence of Selective Social Referencing in Infancy. *Infancy*, (2), 253-264.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28, 675-735.

Triesch, J., Teuscher, C., Deák, G., & Carlson, E. (2006). Gaze following: why (not) learn it? *Developmental Science*, 9(2), 125-147.