

Editorial Manager(tm) for The Review of Philosophy and Psychology  
Manuscript Draft

Manuscript Number: ROPP118

Title: Joint Attention: from Interaction to Joint Action

Article Type: Regular Article

Keywords: joint attention; interaction; joint action

Corresponding Author: Anika Fiebich

Corresponding Author's Institution: Ruhr-University Bochum

First Author: Anika Fiebich

Order of Authors: Anika Fiebich;Shaun Gallagher

Abstract: In this paper, we investigate the role of joint attention in joint actions. We propose a concept of basic joint action analogous to a concept of basic action and argue that joint attention is a basic joint action, which builds the bridge from interaction to joint action. Furthermore, we distinguish between joint path-goal actions and joint final-goal actions depending on the shared intention the agents have.

Suggested Reviewers:

## Authors

Anika Fiebich & Shaun Gallagher

## Title

Joint Attention: from Interaction to Joint Action

## Affiliations and Addresses of the Authors

*Anika Fiebich, M.A.*

Project *Other Minds: Neurophilosophy and Neuroethics of Intersubjectivity*

Ruhr-University Bochum

Carnap Institute of Philosophy and Science

Universitätsstraße 150

44780 Bochum

Germany

Email: [anika.fiebich@rub.de](mailto:anika.fiebich@rub.de)

Tel.: 0049-234-32-24725

Fax: 0049-234-32-14963

*and*

*Shaun Gallagher, Prof. Dr.*

Philosophy and Cognitive Sciences

Institute for Simulation and Training

University of Central Florida (USA)

School of Humanities

University of Hertfordshire (UK)

## Acknowledgments

A. F. acknowledges the support provided by her supervisor Albert Newen and the Barbara-Wengeler-Foundation for research stays at the École Normale Supérieure in Lyon and the Institut Jean Nicod in Paris (2010); S. G. acknowledges the support provided by research grants from the Centre National de la Recherche Scientifique as visiting professor at the École Normale Supérieure de Lyon, and visiting researcher at the Centre de Recherche en Epistémologie Appliquée, École Polytechnique, Paris (2010).

# Joint Attention: from Interaction to Joint Action

## Abstract

In this paper, we investigate the role of joint attention in joint actions. We propose a concept of basic joint action analogous to a concept of basic action and argue that joint attention is a basic joint action, which builds the bridge from interaction to joint action. Furthermore, we distinguish between joint path-goal actions and joint final-goal actions depending on the shared intention the agents have.

Keywords: joint attention, interaction, joint action

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 just pairs of participating agents. Specifically we discuss the role of 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 theoretical account of the role of joint attention in joint actions is still missing in the current literature. The present paper aims to provide the beginning of such an account. We propose a concept of basic joint action analogous to a concept of basic action and argue that joint attention is a basic joint action, which builds a bridge from interaction to joint action. Joint attention is a basic joint action in two ways: (1) it fulfills the minimal conditions to be a joint action, and (2) it is as a constitutive element involved in many but not all complex joint actions. Depending on the shared intention the agents have, we also distinguish between complex joint path-goal actions and joint final-goal actions.

## 1. Interaction and joint action

Following De Jaegher et al. (2010), we take social 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 of the agents is the sole regulator of the coupling whereas the other is just co-present, this is no longer interaction in the strict sense. A 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.

Joint action is a complex form of social interaction that involves a shared intention. 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. On our view, the coordinated behavior patterns are (more or less) specified dependent 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).<sup>1</sup> In addition, *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 a common goal, we are aware that we have a common goal.

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 shared goal in a way that their we-intentions are non-reducible to a set of I-intentions, 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, 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,

---

<sup>1</sup> Carpenter (2009, 381) suggests that the requirement of the shared goal “is missing from [Sebanz et al’s] definition.” It is not clear that Sebanz and colleagues’ definition of joint action lacks the notion of a shared goal, since the phrase *to bring about a change in the environment* indicates such a goal, and the authors even discuss the establishment of such a goal; “for example, when we observe somebody dragging a large table with their hands behind the back we will understand that this person only drags it in such an awkward way to avoid obstacles. Having identified the other’s goal, we can act according to our perception of it, thereby establishing a joint goal. [...] It is likely that similar neurocognitive mechanisms govern goal-directed imitation and the selection of appropriate actions to achieve joint goals” (Sebanz et al 2006, p.71).

even if each has mutual knowledge of the intentions of the other. That's what we call having "a shared goal". It is more than "having a common goal" since mutual knowledge of aiming for that goal is involved here. 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 shared goal but also having a shared goal that involves we-intentions (see table 1). The notion of we-intention implies the notion of cooperation; 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, Searle himself neither uses the term "shared intention" nor makes use of we-intentions as we do in our account. According to Searle, we-intentions are attitudes of single individuals. They are special intending attitudes and need to be distinguished from individual intentions with a special mental content that involves a "we"-activity. Following Bratman (2009), we suggest 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).

Bratman provides his own account of shared intention in which he highlights the idea that a shared intention consists primarily of the interrelations of the agent's attitudes. He (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. Finally, he adds the idea that in joint action we need to mesh sub-plans. 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 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. Although we acknowledge Gilbert's objection 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. Furthermore, we will argue that the meshing of sub-goals may, but do not need to, play a role in joint actions. Accordingly, we bracket them in what we suggest are the conditions for joint action. All three conditions are necessary, and together they are sufficient for joint action:

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

In general, having a shared goal presupposes

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

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

3. participation in cooperative behavior patterns (determined to varying degrees by rules or regularities) in order to achieve the goal [and/or to mesh sub-goals].

Whether and to what extent obligations are involved, or whether meshing sub-goals is required, depends on the shared intention of the particular joint action. The nature of the shared intention involved in a joint action determines how agents need to coordinate their behavior in order to succeed in their cooperation. In this regard we distinguish two different kinds of joint action. In *joint final-goal actions*, the 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 may be achievable in any number of ways. However, agents may also coordinate their behavior as an end in itself; call this a *joint path-goal action*. In this case, the shared intention entails only the activity itself, and thus, the coordinated behavior patterns to achieve the shared goal are specified in advance.

The more rule-governed<sup>2</sup> the shared intention (i.e. the more the coordinated behavior patterns are specified in advance, which is very much the case in joint path-goal actions), the more commitments are involved and the higher are the action expectations and obligations, but the less meshing sub-goals is necessary. 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. But once we accept all the rules entailed in chess, no sub-goals are left that might diverge and stop our joint activity since our action opportunities within the game are prescribed in advance<sup>3</sup>. Additionally, we are not only obliged to observe the rules of chess when we decided to play chess together but we are also committed to finishing the play. 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.

In contrast, if we decide to cook together (a joint final-goal action where the final goal is to have dinner as the end product of our cooking), there is still a lot of space left in this shared goal for diverging sub-goals; e.g., you might think of a rare steak whereas I prefer a vegan dish. Even if we both decide to cook a well-known dish like Spaghetti Bolognese there might

---

<sup>2</sup> In general, one can distinguish between different aspects that are involved in the rules or regularities that determine the coordinated behavior patterns of two agents in order to achieve a shared goal: (1) *natural aspects* (i.e. aspects 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 aspects* (i.e. aspects that are rationally specified such as mathematical or logical rules, scientific methods), and (3) *social/conventional aspects* (i.e. aspects that are specified by conventions like the rules of chess but also social-cultural commitments such as the agreed upon goal of finishing the game).

<sup>3</sup> 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.) the joint activity takes place in.

be divergences in how we go about preparing the dish – I might dislike onions whereas you don't want to leave this ingredient out of the sauce, which might mean that one of us may have to agree to deviate from her original sub-goal so that we don't give up our shared goal. 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. But also joint final-goal actions might (but do not need to) 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 shared intention is, 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 path-goal actions and joint final-goal actions involve the basic joint commitment to pursue the shared intention until its end. In joint final-goal action, this end is achieved when we accomplish the final state or end-product we aimed for (e.g., when we finish a game of chess). In joint path-goal actions, this end is achieved when we bring the action we aimed for to its end, e.g., when we finish our joint action of walking together. In some joint path-goal action however, this end is not specified. If so, no obligations to pursue the shared intention are involved in the joint action. Imagine, e.g. we meet each other in a discotheque. You start to dance with me and I respond to your dancing movements. Here, we are spontaneously committed to dance together. However, no end is specified in this joint commitment. Both of us are in the position to stop dancing to take a drink, or to start dancing with someone else whenever we want, and neither of us are in a position to expect or demand the other to continue dancing to the end of the song.

## 2. 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.

*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) and Leavens and Racine (2009) equate 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, 132), or, “the intentional co-orientation of two or more organisms to the same locus” (Leavens and Racine 2009, 241). In contrast, we suggest that this co-orientation is not yet joint attention, and only becomes *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 an 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.

*Mutual shared attention.* Other theorists define joint attention by adding another characteristic to shared attention, i.e., the *mutual* knowledge of *both* individuals as being attentive towards the same external entity (e.g. Tomasello et al. 2005) accomplished via communicative cues (Carpenter and Liebal, in press). Call this “mutual shared attention”. 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 (Klennert et. al. 1986, Moses et. al. 2001), the infant is engaged not just in shared attention but in *mutual* shared attention, i.e., an interaction that includes the awareness that she and the other are both attentive towards the same entity *and* that the other is aware of this shared attention as well. 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). Communicative cues in social learning situations involving mutual shared attention evoke different learning effects than if those cues were absent; that is supported by empirical evidence from developmental psychology (Csibra 2010).

Usually, mutual shared attention means that individuals are visually mutually attentive towards an external entity. However, mutual shared attention can also include other modalities, e.g. our auditory modality when we enjoy listening to a concert together. Mutual shared 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). Mutual shared attention can also go beyond perception and involve the conceptual. Our conversation about justice, for example, requires that we mutually attend to this concept and the conversation itself confirms that we do.

*Joint attention.* When, as in some of these examples, mutual shared attention involves a shared intention, we have joint attention, in the strict sense. That is, in joint attention, 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 mutual shared attention). On this definition, joint attention fulfills all three conditions of a joint action. In joint attention, there is a shared intention (condition 1), which, at a minimum, may entail the immediate shared goal 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). Joint attention is also involved in complex joint actions that take place in real interactive settings. Here, we need to aim for mutual shared attention in order to cooperate. The coordinated behavior of our cooperation is defined by the shared intention we have (see previous section).



Imagine e.g. we cook together and follow the guidelines of a certain recipe. Here, we aim for mutual shared attention towards our individual actions of what ingredient each of us puts into the cooking pot. Otherwise, we would run the risk of salting the soup twice. We do not aim for mutual shared attention towards our actions as individual actions, but rather, throughout the whole interaction (although we are likely to attend to other things as well) we attend to those actions that determine the success or failure of our cooperation.

Crucially, in joint actions we may be jointly attentive 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 the interaction itself. Whereas the former emerges around the age of one year in ontogeny, the latter emerges much earlier. When a mother goes to pick up her infant, even at a few-months of age the infant will move his body in such a way that he anticipates and cooperates in the pick up; here, mother and baby are engaged in a joint action in which both are jointly attending to the event of picking up.<sup>4</sup> Indeed, in this case, joint attention to the interaction makes the interaction a joint path-goal action.

### 3. Joint attention as a basic joint action

In mutual shared attention the infant is aware of herself and the other being attentive towards an entity or event and aware of the other being aware of that as well. When mutual shared attention is involved, both agents interact in a triadic relation (agent-agent-object/event). Mutual shared 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 joint attention and all three conditions of a joint action are fulfilled. In the minimal case, the shared intention to maintain joint attention is an end in itself. As Aristotle (Nicomachean Ethics, Book I) suggested, some actions are ends in themselves and are done for their own sake, and 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, we usually jointly attend to something for reasons other than mere enjoyment, e.g. for the sake of survival (as a final-goal) or as a constitutive element of action-oriented joint attention in complex joint path-goal action such as having a conversation or dancing the tango.

Given the definition of joint action outlined above, we argue that joint attention is a *basic joint action*, on analogy with what theorists such as Danto (1973), Davidson (2001) and others call *basic actions*. First, we briefly review the literature on the concept of *basic action*, and then define 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

---

<sup>4</sup> Our thanks to Vasudevi Reddy for this example.

to get my drink. Danto (1965), in contrast, appeals to causal relation as the criterion of basicness and emphasizes that “when M performs a basic action, he does nothing first that causes it to happen” (p. 142).

Various examples given of basic actions 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). According to Baier’s analysis (1971, p. 168), each criterion provided in the literature fails to define basic action in all cases.

Chisholm’s criterion, like Danto’s, can then select as basic some actions which are distinguishable from bodily movements, and can fail to discriminate between two actions which must then count as equally basic. Both criteria select conventional and expressive actions as basic, and where Danto’s selects things causally more primitive than bodily movements, namely neuron firings, Chisholm’ avoids this at the cost of selecting things like lace-tying which are causally more complex than bodily movements.

Furthermore, she highlights that actions might be considered basic in various ways; (a) causally (as proposed by Danto), (b) instrumentally (as proposed by Chisholm), or (c) conventionally most primitive, (d) basic in a compositional, (e) logical, or (f) genetic sense, (g) basic dependent on how easy it is to perform the action, and (h) how easy in isolation from the others. Such criteria, moreover, may be incompatible with each other; e.g. whereas an action may be basic in a genetic sense, it does not need to be the easiest to perform.

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 a 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)

The controversies surrounding basic actions, however, do not apply to the concept of a basic *joint* action. There is no claim that we perform a basic joint action without performing some other non-joint actions, and, unlike the concept of basic action, there is a clear causal (or motivational) history that can be defined for a basic joint action such as joint attention. Whereas the debate about basic action concerns whether certain bodily events constitute actions or not, and whether these events are necessary for and logically prior to more complex action, basic joint actions are *joint* actions defined by clear conditions involving (at least) two interacting agents and their *shared* intention (see definition of joint action above).

We thus argue that joint attention is a basic joint action. 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. Indeed, one can analyze joint attention into more subtle components of interaction (those that involve mutual 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. Since joint attention, insofar as it involves (1) a shared intention involving we-intentions, (2) common knowledge of intending a shared 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, then it is a basic joint action, and in many (and perhaps most) cases of joint action, it operates as such.

Whereas mutual shared attention may be involved in triadic interactions, joint attention (i.e. mutual shared attention for a mutually acknowledged purpose) is, as we will show, involved in joint actions. Since joint attention is a basic joint action, it builds the bridge from interaction to joint action. Joint attention as a basic joint action involves the shared intention to engage in mutual shared attention either as a means to some other end, or as an end in itself. Joint attention is a (shared) intentional phenomenon that can play a role in complex joint actions that take place in real interaction settings, and without joint attention these more complex actions would not happen. In those cases, joint attention serves purposes other than itself, e.g., 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.

We made a distinction between two different kinds of joint actions, depending on the shared intention involved: (1) *joint path-goal action*, and (2) *joint final-goal action*. Based on the above analysis we suggest that joint attention as a basic joint action is included in all complex joint path-goal actions but only in those complex joint final-goal actions that take place in real interactive settings. We make this clear in the following section.

#### 4. Joint attention in joint action

Joint attention as a basic joint action is included in a number of complex joint actions that take place in real interactive settings, but in different ways, depending on the shared intention and the kind of joint action involved.

Let's start with an example of an action that involves a minimal 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 path-goal action. We intend to walk together even though we have different final goals – literally, final destinations. We adjust our gait in order to stay together as we walk. After that there is minimal perceptual monitoring of our walking action – so we don't bump into each other or trip each other as we walk. Likely we engage in conversation

as we go, considering that the joint action of walking together is not simply about moving our bodies. While we may not attend explicitly to our gait, we do attend to each other and to various things as we walk and talk, in a way that involves joint attention and that keeps both the walking and the conversation going; e.g. we avoid obstacles in the road in a coordinated fashion; perhaps by implicit agreement we avoid sensitive topics (e.g., the recent labor unrest among clowns).

In some joint path-goal actions, as in the example just mentioned, the coordinated behavior pattern just is the action as an end in itself, and we maintain joint attention to avoid the risk that our shared intention to walk together fails; e.g. because of obstacles in the road. We cooperate and coordinate our behavior in a way that is determined by the natural aspects of our bodies and the environment. Joint attention may also serve to improve performance, e.g. for two professional dancers doing the tango. In more complex forms of joint path-goal actions, stronger conventional commitments may be involved, e.g., in playing a conventional game like chess. In the latter case, we jointly attend to each move and joint attention is involved in almost the whole interaction. Here also, the shared goal involving We-intentions (to play chess) and the cooperative behavior patterns to achieve this goal (playing chess) are the same. Whether there is a final end product or state included in the shared intention or not (me winning the game is not a shared intention, but that someone wins may be part of the shared intention), it is worth mentioning that the shared intention includes performing the action *to its end*, i.e., the commitment to finish the joint action regardless of any sub-goals. Clearly, however, we need to pay close and explicit attention to the action itself – both of us jointly attending to our individual moves on the chessboard.

In joint final-goal actions, the shared intention entails a final end-product or state. If we decide to do a puzzle together, we aim for a certain final product of our joint activity, namely for creating a completed puzzle. Even in such a joint final-goal action, we might enjoy the cooperative behavioral patterns that are required for achieving the final product so much that we may partly have as our shared intention doing the puzzle because of the activity itself. But we could also dislike the activity itself and aim nonetheless to finish the puzzle. Joint final-goal actions are independent of whether the shared goal involving We-intentions involves the sub-goal of enjoying the activity itself. Rather they are defined by the final product for which the agents aim. But in doing the puzzle, joint attention is involved in an extremely high degree due to the constraints of the task, even if in some instances we are working independently on different corners of the puzzle.

In some joint final-goal actions, joint attention is required to keep cooperative behavior patterns on track to achieve the goal, e.g., if we cook together and you see me salt the soup, you wouldn't salt the soup again. In other joint final-goal actions however, joint attention is not necessary. 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 mutual shared 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 joint attention that in other cases might be required in real interactive settings to succeed in achieving the shared goal. Nonetheless, the three defining conditions of joint action are fulfilled.

In joint actions there are matters of degree and a variety of ambiguities concerning the formation of shared intentions, the role of joint attention, 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 and from mutual shared attention due to the environmental circumstances of the social interaction or the intentional behavior of one of the agents.

In some situations if the forming 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 shared goal *together*. 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.” 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 as a basic joint action is frequently required, or some set of instructions 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.

## 5. Summary and conclusion

In this paper, we discussed the role of joint attention in joint actions. Analogous to a concept of basic action, we proposed that joint attention is a basic joint action and argued that considered in this way, joint attention builds a bridge from interactions to joint actions. We distinguished between two kinds of joint actions; i.e. joint path-goal actions and joint final-goal actions. Joint attention is involved in joint path-goal actions, for example, to maintain coordination, to improve performance, or when agents run the risk of failing to achieve their shared goal. That can also be the case for joint final-goal actions, but here, other mediators might serve to ensure success.

## 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 Jaegher, 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.
- Leavens, D. A., & Racine, T. P. (2009). Joint attention in apes and humans. Are humans unique? *Journal of Consciousness Studies*, 16, 240-267.
- 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. *Cogniție, 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.

Table 1: Goals and intentions

<i>Common goal</i>	Two or more agents having the same goal (without common knowledge of aiming for the same goal)
<i>Shared goal</i>	Two or more agents having the same goal plus common knowledge of aiming for the same goal
<i>We-intention</i>	An individual’s special intending attitude to cooperate toward attaining a shared goal
<i>Shared intention</i>	Two or more agents having a shared goal involving we-intentions