9

A JOINT ACTION PERSPECTIVE ON EMBODIMENT

Robrecht P.R.D. van der Wel, Natalie Sebanz, Günther Knoblich AuQ71

The field of joint action research has rapidly emerged from the realization that studying the mind exclusively in insular contexts may be insufficient for fully understanding how cognition works (Sebanz, Bekkering, & Knoblich, 2006). To go further, one may argue that several cognitive functions are shaped, and in some cases only exist, to engage in joint contexts. Language provides a clear example in this regard, as one could imagine the hypothetical scenario of a world in which individuals would not engage in any social interaction. In such a scenario, the functionality of knowing a language would be rather minimal, and language may not have existed under such a pretense in the first place.

People interact in other ways than just through language, however. In the emerging joint action literature, it is indeed often noted that people can build bridges, carry furniture, and dance the tango together. Whereas some of these joint actions, such as finger tapping, dancing or singing together, create social connection (e.g. Hove & Risen, 2009; Wiltermuth & Heath, 2009) and may constitute expressions of group membership and culture, other joint actions, such as building bridges, may be driven by the limits of our own bodies. A single individual simply could not have built the Brooklyn Bridge, the Taj Mahal, or the house you may live in, indicating that the limits of our own bodies encourage joint action. By planning and coordinating joint actions, people have managed to build the structures and infrastructures that make our world what it is today.

How are joint actions achieved, and in which ways do they relate to core tenets of embodiment? In this chapter, we will demonstrate that successful joint action performance relies on sophisticated cognitive and sensorimotor mechanisms that are shaped by the action abilities of our bodies. As we will show, an enhanced understanding of joint action mechanisms can help to constrain theories of embodiment as it reveals the relative contributions of situated online processing and symbolic thinking to successful joint action performance.

We have structured this chapter in two parts. First, we will briefly review some of the main tenets of embodiment that have been proposed. We will examine how the case of joint action speaks to these tenets. In the second part, we will then go on to review different aspects of joint actions, and what is known about the mechanisms contributing to joint action performances. In particular, we will introduce and substantiate specific mechanisms that support emergent coordination. Some of these mechanisms are body based, whereas others rely strongly on the environment. Then, we will review evidence on how planned coordination comes about. We will give examples of how these various sub-processes together interact to give rise to the wealth of joint action performances we encounter every day. We will end by discussing the sense of agency over joint performances, and will show that an embodied approach is warranted in that domain as well.

Six tenets of embodiment

In a foundational article, Wilson (2002) synthesized six tenets to which proponents of embodiment may subscribe. These tenets were that 1) cognition is situated; 2) cognition is time-pressured; 3) we off-load cognitive work onto the environment; 4) the environment is part of the cognitive system; 5) cognition is for action; and 6) offline cognition is body based. At the time of Wilson's analysis, researchers had hardly begun to study joint actions experimentally. As joint action research has emerged as an active research domain within the cognitive sciences and related fields in the past decade or so, we will discuss how these tenets link to joint action research in turn.

Cognition is situated

As we already indicated, the field of joint action research emerged in part because researchers realized that much of cognition is for interacting with others. To understand cognition, experimental approaches until then relied mostly on individual participants performing some (often fairly arbitrary) task by themselves. Although these approaches yielded substantial knowledge about cognition, studying the mind situated in joint action contexts has for example shown that cognition in seemingly simple stimulus-response paradigms is influenced by the presence of co-actors (e.g. Sebanz, Knoblich, & Prinz, 2003). Thus, joint action research aims to understand cognition from a situated perspective. We will discuss this issue more when we consider planned joint coordination.

Cognition is time pressured

Coordinated joint actions by definition require precise timing of the actions of multiple actors to achieve a common goal. Indeed, joint action has been very broadly defined as a social interaction whereby two or more individuals coordinate their actions in space and time to bring about a change in the environment

(Sebanz, Bekkering, & Knoblich, 2006). For example two actors who want to lift the two ends of a tray full of glasses together, need to precisely coordinate to initiate their actions at the same time and to adjust the applied forces in real time to maintain balance. If either of the actors fails to do this the joint action will fail. Thus, the cognition underlying successful joint actions is often inherently time pressured due to the need to collaborate and coordinate with others.

The time-pressured nature of cognition makes reliance on mechanisms that obviate the need for slow and effortful computations to support cognition particularly beneficial. When lifting a table, for example the use of online perceptual and motor information could be put to service to accomplish the task. Rather than representing the underlying mechanics of table lifting and the influence of each actor on such mechanics, actors could rely on the haptic information they get from the movements of the table itself once it is lifted. Reliance on slow cognitive mechanisms would be much less efficient than relying on embodied mechanisms.

Off-loading cognition onto the environment

Although philosophers generally agree that joint actions involve intentions that are shared among actors, there has been considerable disagreement about what such shared intentions entail. For example some have argued that shared intentions encompass more than the sum of individual intentions (Gilbert, 1992), but others hold that joint actions rely on the meshing of individual sub-plans (Bratman, 1992, 2009; see further Tollefsen, 2005). In addition, some have viewed shared intentions to involve detailed representations of co-actors' tasks (Bratman, 1992; Tomasello, Carpenter, Call, Behne, & Moll, 2005), whereas others have started with minimal representational requirements (Clark, 1997; Vesper, Butterfill, Knoblich, & Sebanz, 2010). From a minimalistic representational view, one could consider reliance on other actors to accomplish a shared goal as a form of cognitive off-loading. By forming a shared goal, it is not necessary to fully represent all aspects of to be completed tasks. Instead, one could employ one's cognitive processes to accomplish one's own contribution to the task, and represent only those aspects of the overall task and of the co-actors' contributions that are needed for successful coordination (see Vesper *et al.*, 2010, for a minimal architecture of joint action). Such an off-loading may be especially pronounced when two or more actors differ in their expertise with performing different parts of a task (Wegner, 1987).

The environment as part of the cognitive system

If one follows the strong version of the distributed cognition claim, then minds and their surrounding environments ought to be considered together, as parts of the same system (Hutchins, 1995). For instance, the pilots in a cockpit, the cultural artifacts they operate (the plane with its instruments and specific build), and the physical laws governing flight may all need to be considered together to understand the cognitive processes at play. A related idea in the motor domain is that of interpersonal synergy where the movements of two actors become so tightly coupled that they are best conceived of as a single system rather than as two separate individuals (Ramenzoni *et al*., 2011; Schmidt, Carello, & Turvey, 1990). While such views have initiated and continue to inspire research on joint action many joint action researchers follow a weaker version of the distributed cognition view where individual cognition is still considered the main target of explanation. In this view the question is how our minds are designed for and shaped by interactions with other agents.

Cognition is for action

Whereas it is clear that joint action researchers value the importance of studying action for studying cognition, doing so does not necessitate subscription to the claim that cognition is for action in a strict sense. Nonetheless, joint action research has provided evidence for a tight link between cognition and action. For example it has been demonstrated that anticipating reaching and grasping actions of a joint action partner results in similar preparatory activation of the motor system as when one is preparing to perform the same actions oneself. The anticipatory motor activation does not occur when anticipating actions by people who do not take part in the joint action (Kourtis, Sebanz, & Knoblich, 2010). Further studies show that preparation for the partner's actions is driven by processes that reflect joint action planning (Kourtis *et al*., in press; Kourtis, Sebanz, & Knoblich, 2013). Participants' own overt actions have also been shown to be affected by the beliefs of others, even when such beliefs were task-irrelevant (e.g. van der Wel, Sebanz, & Knoblich, 2014). Such findings suggest a clear link between cognition and action, but they do not imply that cognition is specifically for action, or that cognition and action always link to one another directly (see Wilson, 2002).

Offline cognition is body based

Wilson (2002) summarized support for the claim that offline cognition is body based by providing examples from a wide range of cognitive functions, including mental imagery, several memory sub-systems, and reasoning and problem-solving. As discussed throughout this book, evidence for this claim has only increased in recent years. With respect to joint action research, perhaps the most relevant demonstrations of the use of the motor system for offline cognition have come from studies on action observation. Indeed, it has been shown that one's own motor system is active when observing others' actions (e.g. Buccino *et al*., 2001; Cross, Hamilton, & Grafton, 2006; Cross, Kraemer, Hamilton, Kelly, & Grafton, 2009) and when imagining another's actions (Grèzes & Decety, 2001; Ramnani & Miall, 2004). Such motor resonance has also been shown to depend on the familiarity with the action (Casile & Giese, 2006; Knoblich & Flach, 2001), on one's own motor expertise (e.g. Aglioti, Cesari, Romani, & Urgesi, 2008; Calvo-Merino, Glaser, Grèzes, Passingham, & Haggard, 2005) and, as we already mentioned, on the social relationship between the observer and the actor (Kourtis, Sebanz, & Knoblich, 2010).

Summary

The preceding analysis indicated a clear link between joint action research and the central tenets of embodiment. The field of joint action research takes a situated approach, acknowledges the importance of timing, and investigates whether off-loading happens and how offline cognition uses our body. By emphasizing the importance of actions, joint action research is amenable to the possibility that cognition is for action. In the remainder of this chapter, we will further lay out what is known about the mechanisms that support joint actions.

Emergent and planned coordination

Joint action is an umbrella term that comprises a wide range of different actions. Broadly, we will conceptualize joint actions to concern two types of coordination, planned coordination, and emergent coordination. These types of coordination differ in the extent to which they rely on intentionality and representations that specify the desired outcomes of joint actions.

Planned coordination is intentional in nature and requires some form of representation of the goal of the joint action as well as the actor's own contribution to achieve the desired joint action outcome. Below, we will discuss evidence that people tend to automatically represent a co-actor's task (Sebanz *et al.*, 2006), but how much detail such representations contain may vary greatly. In some cases, others' motives, thoughts, or perspectives may be taken into account, whereas in other cases people may represent others at a minimal level and simply wait for a particular action to happen (Vesper *et al.,* 2010). In emergent coordination, coordinated behavior occurs relatively automatically due to for example perception-action couplings that make multiple individuals act in similar ways.

Emergent coordination does not rely on joint plans or require common knowledge. Instead, agents may process the same perceptual and motor cues in similar ways, resulting in spontaneous coordination. For example when two people start walking when a traffic light turns green, they share the timing of the perceptual event of the light turning green. In addition, people who walk next to one another tend to synchronize their behavior (van Ulzen, Lamoth, Daffertshofer, Semin, & Beek, 2008). Thus, multiple agents may seemingly start to act as a single coordinated entity (Marsh, Richardson, & Schmidt, 2009; Spivey, 2007) as the same cues and motor routines drive behavior in the involved individuals. As this example indicates, emergent coordination is highly situated. We will discuss how emergent coordination arises next.

Emergent coordination and joint action

Emergent coordination occurs in many physical systems, and is not restricted to biological systems. Christiaan Huygens (1673/1986) first observed that two clocks hanging on the same wall will tend to fall in synchrony with one another, purely due to their physical coupling. Thus, for emergent coordination to occur does not require intentions, a brain, or even a nervous system. As humans are physical systems, it is not surprising that our behaviors display emergent coordination as well. Psychologists who have studied behavior from a dynamical systems perspective have indeed done so successfully by characterizing a wide range of behavior to arise from self-organizing coupled oscillator models (Haken, Kelso, & Bunz, 1986). AuQ72

Several sources may give rise to or contribute to emergent coordination. These are entrainment, common affordances, and perception-action matching. We will consider each of those in turn, as well as their contribution to joint action.

Entrainment

Entrainment refers to the tendency for spatiotemporal coordination to occur spontaneously between two parts of a moving system. These parts are not necessarily directly linked, implying that entrainment may happen between two people (Schmidt & Richardson, 2008). Indeed, studies on human movement coordination have provided evidence for entrainment in interpersonal settings. In several of those studies, pairs of participants were instructed to swing pendulums or legs alongside one another, while maintaining their preferred tempo. When they could see each other, participants tended to entrain, such that they moved in synchrony more often than would be expected by chance (Schmidt & O'Brien, 1997). This was even the case when participants rocked in rocking chairs that had different natural frequencies (Richardson, Marsh, Isenhower, Goodman, & Schmidt, 2007). It has similarly been shown that people entrain when they engage in conversation by synchronizing their body sway (Shockley, Santana, & Fowler, 2003) and when audiences in Romania and Hungary clap in unison for a theatre or opera performance (Neda, Ravasz, Brechte, Vicsek, & Barabasi, 2000).

Entrainment is thought to concern relatively low-level cognitive activity, as it does not rely on intentions or action goals. With regard to embodiment, entrainment forms a clear example of how cognition is situated and of how behavior arises from the interaction between the body and the environment.

Affordances

The term 'affordance' refers to the action opportunities objects and the environment in general provide to an agent with a particular action repertoire (Gibson, 1977). For example chairs afford sitting, cups afford grasping, and flat, even terrains afford walking and biking. Because different people have similar action repertoires and may perceive the same objects, they share common affordances. Such affordances form another source for emergent coordination when multiple agents perceive the same environment and objects at the same time, as it makes it likely that the involved agents perform similar actions. Thus, when people are sitting in the grass and it starts to rain, they may simultaneously run towards a gazebo for shelter (Searle, 1990), as the shelter has the affordance to keep them dry. In such a case, the involved agents do not need to intend to coordinate with one another, but coordination emerges.

Interestingly, when people work together affordances emerge that may not be present for an individual. For example a large or heavy object may afford lifting by two or more people, but not by a single individual. We will refer to such affordances as joint affordances (Knoblich, Butterfill, & Sebanz, 2011). It is likely that joint affordances often result from a combination of planned and emergent coordination. For example when people need to move planks of varying lengths from a conveyor belt, they may lift short planks individually but longer planks together. When they switch from one mode to the other may depends on the relationship between the plank's length and the pair's joint arm span (Richardson, Marsh, & Baron, 2007).

As this plank-lifting example indicates, affordances are not necessarily restricted to the bodies of single individuals, but may arise through the embodied characteristics of joint actors. The ability to perceive and act upon joint affordances provides a much wider range of action opportunities for joint action.

Perception-action matching: Common mechanisms

Traditionally, cognitive scientists considered cognition to involve symbolic codes that were akin to the operations of a computer. Within that view, the motor system only provided a mechanism to translate the symbolic codes into physical actions (e.g. Anderson & Bower, 1973, Marr, 1982; Newell, Shaw, & Simon, 1958; Kladzky, 1975, Kieras & Meyer, 1997). In contrast to this approach, proponents of embodiment have argued that cognition is for action, and actions fundamentally shape cognition itself. Throughout this book, much evidence consistent with this notion (or at least that actions shape cognition) has already been provided. For example the action system has been shown to interact with the language system, with the formation of conceptual knowledge, and with object perception.

How does the link between cognition and action support emergent coordination? The answer to this question specifically concerns evidence indicating that action perception and action production rely on common mechanisms (see van der Wel, Sebanz, & Knoblich, 2013, for a recent review). This possibility is at the core of the common coding theory (Hommel, Muesseler, Aschersleben, & Prinz, 2001; Jeannerod, 1999; Prinz, 1997), which formed an extension of William James's (1890) ideomotor theory of voluntary action. These codes do not represent actions per se, but rather their distal perceptual effects. If perception and action rely on common codes, this makes the integration of one's own and co-actors' action effects for joint actions relatively straightforward. It also implies that considering perception and action in a joint context is fruitful based on an embodied perspective.

Consistent with common coding, single-cell studies in monkeys and brain imaging studies in humans have found similar activation patterns during action production and action observation (Rizzolatti & Sinigaglia, 2010). In monkeys, this evidence is especially strong, as the discovery of mirror neurons indicates a mapping between observation and execution in single neurons. When people observe others' actions, the amount of activity in the motor system also depends on how well the observer's own action repertoire maps onto the observed actions (Calvo-Merino, Glaser, Grèzes, Passingham & Haggard, 2005; Calvo-Merino, Grèzes, Glaser, Passingham, & Haggard, 2006; Cross, Hamilton, & Grafton, 2006; Cross, Kraemer, Hamilton, Kelly, & Grafton, 2009). Single-neuron evidence in humans has also been obtained from patients implanted with intra-cranial depth electrodes to identify the loci of seizures (Mukamel, Ekstrom, Kaplan, Iacoboni, & Fried (2010).

As observed actions are matched onto the observer's own action repertoire, the resulting activity in the motor system increases the likelihood for an actor to produce the observed action (resulting in mimicry, Chartrand & Bargh, 1999). Thus, perception-action matching contributes to emergent coordination, and it may do so interpersonally when multiple agents observe the same action at the same time.

Aside from modulating activity in the motor system during action observation, perception-action matching also supports action simulation. The reason for this is that the same internal predictive models that are active during action production (Wolpert, Doya, & Kawato, 2003) may be employed to predict another agent's actions in real time. Perception-action matching may support joint action by providing a clear interface for coordinating one's own actions with those of others. As action production and action observation rely on the same codes, integrating the codes for a co-actor with the codes for one's own action contribution is relatively easy. We will see at the end of this chapter that this integration poses a challenge for understanding how people derive a sense of agency over joint actions, however.

Although emergent coordination may occur without an intention to coordinate and although some coordination processes of emergent coordination may also operate outside of the domain of social interactions, such processes can be crucial in enabling coordination in goal-directed joint actions. However emergent coordination is not the only way to achieve coordination during joint action. We will now discuss additional mechanisms under the header of planned coordination. Such coordination depends on representing the outcomes of joint actions and individuals' contributions to them. The discussed mechanisms provide an interesting challenge for theories of embodiment, as it is not readily apparent how such theories account for them.

Planned coordination and joint action

People often intentionally plan to achieve a common goal with one or more other agents. In some cases, such joint actions may only require that the other agents' actions are represented at a minimal level. One's own contribution and the overall goal need to be represented, but the identity of other agents and their contributions to the joint action do not need to be. Vesper *et al.* (2010) captured such minimal representational requirements with the formula 'ME + X'. In this formula, X refers to anything outside of one's own contribution that contributes to the joint action. For example when people do the dishes together, the person drying the dishes would only need to wait until a washed plate is put into the dish rack, rather than needing to represent the actions of the other agent who is doing the actual washing. Similarly, when a violinist plays in an orchestra, the involved task representation may concern one's own contribution to the symphony, as well as a representation of the overall sound, but not the details of each of the other individuals' contributions. Thus, minimal representations suffice in such cases.

In other cases, planned coordination may rely on more extensive representations of the other agents in the joint action. For example when two people lift an object together, it is important that each actor represents where and when the other actor is grasping and lifting it. When people hand each other objects, it has been shown that they form more extensive representations of the co-actor's task over time (Meyer, van der Wel, & Hunnius, 2013). In the domain of language, when two or more agents engage in a conversation, each agent likely tries to form a representation of the other agents' views and intentions in order to understand them.

We will now consider two processes that contribute to planned joint coordination, and the evidence for them. These processes are the formation of shared task representations and joint perceptions.

Shared-task representations

The field of joint action research has rapidly grown over the past decade, due in part to the discovery of shared-task representations. Evidence for such representations originally came from an experiment in which two people performed a classic Simon task (Simon, 1969) together instead of by themselves (Sebanz, Knoblich, & Prinz, 2003, 2005). In the individual version of the task, a participant sat behind a monitor and viewed an index finger with a colored ring placed on it. The ring could either be green or red, and the participant was asked to indicate the color of the ring by pressing the left button for one color and the right button for the other color. Importantly, aside from variations in the ring's color, the stimuli also differed in terms of the pointing direction of the index finger. Although this feature was task irrelevant, participants showed a standard Simon effect, as they responded faster when the location of the required button press to identify the ring color happened to correspond with the pointing direction of the index finger versus when it did not. The interesting part of the experiment concerned the joint task condition, in which two participants each did half of this task. Thus, one participant only pressed the response button on the right if the ring was their assigned color (say, red), and the other participant pressed the button on the left if the ring was the other color (say, green). Thus, each participant effectively performed a standard go–no go task.

The findings indicated that participants showed a Simon effect in the joint condition. They did not show this effect when they performed half of the task by themselves, without another actor taking care of the other response button. Thus, the results suggested that people automatically co-represent a co-actor's task as if it were their own. From these results, the existence of shared-task representations was postulated. Although the exact reasons for the joint Simon effect have been debated (e.g. Dolk, Hommel, Prinz, & Liepelt, 2013) and boundary conditions exist (e.g. Guagnano, Rusconi, & Umiltà, 2010), other experimental tasks that employed a similar logic corroborate these findings (e.g. Atmaca, Sebanz, Prinz, & Knoblich, 2008, Atmaca, Sebanz, & Knoblich, 2011; Bates, Patell, & Liddle, 2005; Baus *et al*., 2014; Böckler, Knoblich, & Sebanz, 2012; Eskenazi, Doerrfeld, Logan, Knoblich, & Sebanz, 2013; Heed, Habets, Sebanz, & Knoblich, 2010; Ramnani & Miall, 2004; Schuch & Tipper, 2007; van Schie, Mars, Coles, & Bekkering, 2004). Shared-task representations support planned coordination by allowing actors to know and anticipate what co-actors will contribute to a joint action.

It is important here to elaborate on differences between perception-action links that feed into emergent coordination, and the notion of shared co-representations. Although observation of an action may invoke shared-task representations, the notion of shared-task representations is broader than perception-action links. For example, shared-task representations also become activated when two people perform a memory task together that does not involve any action during the encoding phase. In that case, participants remembered more of their task partner's category than of an unassigned word category when they were given a surprise memory task. This effect even occurred when participants were paid to remember as many words from their own category as possible (Eskenazi *et al*., 2013). Thus, shared-task representations are broader than just action observation contexts.

Joint perceptions

Planned coordination may also benefit from the inclusion of another agent's perceptions into one's own representation of the other's task. For example representing what co-actors in a joint action can see in the environment (Brennan & Hanna, 2009) or what they are looking at (Brennan, Chen, Dickinson, Neider, & Zelinsky, 2007) may help in the coordination with those co-actors. Some have argued however that perspective taking is relatively slow and cognitively demanding, and therefore of limited use for real-time coordination tasks (Shintel & Keysar, 2009).

Several recent studies suggest that perspective taking may be more automatic than originally thought. For example Samson and colleagues (Samson, Apperly, Braithwaite, Andrews, & Bodley Scott, 2010) showed in a visual perspective-taking paradigm that participants took an avatar's perspective, even when that perspective was never task-relevant. We recently obtained similar findings in a belief tracking task (van der Wel, Sebanz, & Knoblich, 2014) where an onlooker's irrelevant beliefs about object locations influenced people's reaching movements. These findings suggest that one's own and others' perspective-taking may happen automatically and in parallel (see also Ramsey, Hansen, Apperly, & Samson, 2013).

Emergent and planned coordination interact during joint actions

We have already provided some examples of how emergent and planned coordination may interact to support joint actions. Speaking to the link between planned coordination and affordances, Richardson *et al.* (2007) found that the shift from individual to joint performance in a plank-lifting task depended on the ratio of the plank's length to the action partners' joint arm span.

There have also been studies linking planned joint coordination and entrainment. In one such study (van der Wel, Sebanz, & Knoblich, 2011), participants learned a new coordination task either alone (bimanually) or together with an action partner (unimanually). The task involved moving a pole (resembling a pendulum) back and forth between two targets by pulling on two strings (one on each side) at the base of the pole. The results indicated that individuals and dyads learned this coordination task at similar rates. Importantly, dyads entrained more than individuals did, as evidenced by the increase in overlapping forces exerted on the two sides of the pole. Generating such force overlap supported emergent coordination by providing haptic information about the action partner, thereby reducing the need to represent the other actor's actions.

Planned joint coordination may also benefit from action simulation. In a recent study, we (Vesper, van der Wel, Knoblich, & Sebanz, 2013) tested whether dyads are able to coordinate when they know each other's respective task, but do not have any real-time information available about their partner's performance. In particular, participants made jumping movements side by side to targets that varied in distance. They could not see their action partner, but received information about their own jump target and their partner's jump target through sets of lights on the floor. After they heard a tone, their task was to try to land on the targets as synchronously as possible. The results indicated that dyads could perform this task surprisingly well, and did so by running a motor simulation of their own and their action partner's jumps. Specifically, the partner with the shorter jump modified both their onset time and their jump height (suggesting motor simulation), depending on the difference in distance between their own and their partner's jump. In addition, both actors started their jumps closer to the start signal than they did in the individual control conditions. Thus, dyads exploited the timing of the shared auditory start signal to align their actions. In another study, we similarly found that dyads used such speeding as a strategy for planned coordination (Vesper, van der Wel, Sebanz, & Knoblich, 2011).

Aside from bottom-up information supporting planned coordination, top-down information has also been shown to influence entrainment in communicative settings. For example Richardson and colleagues (Richardson, Dale, & Kirkham, 2007) asked dyads to converse (a form of planned coordination) about a Dali painting after they had received either the same or different background information about Dali's art. These authors found that eye movements were more entrained for dyads receiving the same information that for dyads receiving different information. Thus, common knowledge influenced the extent to which emergent coordination occurred. In a related study, Richardson and Dale (2005) also showed that listeners better comprehended a monologue when their gaze was more coordinated with the gaze of the speaker. In line with claims of embodiment, such findings indicate a tight link between the environment, low-level sensorimotor activity, and higher-level cognition.

Joint agency and embodiment

Our overview of mechanisms that contribute to successful joint action performances indicates that others' actions influence people's own actions, both for emergent and planned coordination. In line with claims of embodiment, much evidence indicates that perceiving others' actions relies on similar mechanisms as performing those actions, that similar entrainment to environmental input may happen when multiple actors are in the same situation, and that people have a tendency to co-represent others' tasks. Although these mechanisms support integration of one's own and other actors' actions into a joint action performance, it raises a challenging question with respect to how actors distinguish between their own and others' contributions to the joint action. Said differently, how do actors in a joint action derive a sense of agency over such actions? This question is inherently embodied, as it concerns how a higher-level reflective process depends on perceptual and motor information in a situated, time-pressured setting.

Many studies have investigated the sense of agency in individual task contexts, by manipulating the presence of alternative causes for an action and by introducing noise in the performance (see van der Wel & Knoblich, 2013, for a review). How people derive a sense of agency when they intentionally perform an action jointly has only recently started to be investigated experimentally. In one study (van der Wel, Sebanz, & Knoblich, 2012) that employed the pole paradigm discussed above, it was found that the sense of agency of actors in a dyads strongly linked to the objective quality of performance, as has been found in studies on individual agency as well (e.g. Metcalfe & Greene, 2007). Interestingly, the individual forces participants generated (which is a proxy for their actual contribution) did not correlate with the sense of agency in this task.

Dewey, Pacherie, and Knoblich (2014) recently elaborated on how an actor's individual contribution may increase the sense of agency for joint actions. In their studies, participants controlled the movements of a dot while tracking a moving target on a computer screen, and did so either by themselves or together with an action partner. Participants in this study were sensitive to their own contributions when such contributions were clearly distinguishable. When the movements of both actors could have similar perceptual consequences, the sense of agency decreased, however. Thus, the sense of agency for joint actions seems highly situated, and arises from a combination of perceptual and sensorimotor information, as well as causal task structure. Thus, an embodied approach to the sense of agency for joint actions promises to be a fruitful avenue for further research on this topic.

Conclusions

We started this chapter by examining how joint action research links to the main tenets of embodiment. Joint action research generally takes a situated approach by considering cognitive activity in social contexts. It takes actions as a starting point for understanding cognition, and does so in time-pressured contexts. Our overview of emergent and planned coordination indicates that some instances of joint action may entail the close interaction between perception, action, and the environment that has been postulated by radical embodiment approaches. Other forms of joint action, however, also require anticipation, planning, and thinking in order to guarantee successful coordination. The evidence for shared-task representations and joint perceptions appears to provide a challenge for theories of embodiment, as it is not readily apparent how for example joint memory effects would come about based on embodied processing. Future studies of joint action should clarify in which sense the different cognitive processes supporting joint action are embodied, how emergent and planned processes are integrated, and how they affect the phenomenology of individuals acting together.

References

Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience, 11*, 1109–1116.

Atmaca, S., Sebanz, N., & Knoblich, G. (2011). The joint flanker effect: Sharing tasks with real and imagined co-actors. *Experimental Brain Research, 211*, 371–385.

Atmaca, S., Sebanz, N., Prinz, W., & Knoblich, G. (2008). Action co-representation: The joint SNARC effect. *Social Neuroscience, 3*, 410–420.

Bates, A. T., Patel, T. P., & Liddle, P. F. (2005). External behavior monitoring mirrors internal behavior monitoring: Error-related negativity for observed actions. *Journal of Psychophysiology, 19*(4), 281–288.

Baus, C., Sebanz, N., de la Fuente, V., Branzi, F. M., Martin, C., & Costa, A. (2014). On predicting others' words: Electrophysiological evidence of prediction in speech production. *Cognition, 133*, 395–407.

Böckler, A., Knoblich, G., & Sebanz, N. (2012). Effects of co-actor's focus of attention on task performance. *Journal of Experimental Psychology: Human Perception and Performance, 38*(6), 1404–1415.

Bratman, M. (1992). Shared cooperative activity. *The Philosophical Review*, *101*(2), 327–341.

Bratman, M. (2009). Modest sociality and the distinctiveness of intention. *Philosophical Studies*, *144*(1), 149–165.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., & Freund, H. J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience, 13*(2), 400–404.

Calvo-Merino, B., Glaser, D. E., Grezes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex, 15*, 1243–1249.

Calvo-Merino, B., Grèzes, J., Glaser, D. E., Passingham, R. E., & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology*, *16*, 1905–1910.

Casile, A., & Giese, M. A. (2006). Non-visual motor learning influences the recognition of biological motion, *Current Biology*, *16*(1), 69–74.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893–910.

Clark, A. (1997). *Being there: Putting brain, body, and world together again* Cambridge, MA: MIT Press.

Cross, E. S., Hamilton, A. F., & Grafton, S. T. (2006). Building a motor simulation de novo: Observation of dance by dancers. *NeuroImage*, *31*(3), 1257–1267.

Cross, E. S., Kraemer, D.J.M., Hamilton, A. F., Kelley, W. M., & Grafton, S. T. (2009). Sensitivity of the action observation network to physical and observational learning. *Cerebral Cortex, 19*(2), 315–326.

Dewey, J. A., Pacherie, E., & Knoblich, G. (2014). The phenomenology of controlling a moving object with another person. *Cognition*, *132*, 383–397.

Dolk, T., Hommel, B., Prinz, W., & Liepelt, R. (2013). The (not so) Social Simon effect: A referential coding account.*Journal of Experimental Psychology: Human Perception and Performance*, *39*(5), 1248–1260.

Eskenazi, T., Doerrfeld, A., Logan, G. D., Knoblich, G., & Sebanz, N. (2013). Your words are my words: Effects of acting together on encoding. *Quarterly Journal of Experimental Psychology, 66*, 1026–1034.

Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: Towards a general ecological psychology* (pp. 67–82). Hillsdale, NJ: Erlbaum.

Gilbert, M. (1992). *On social facts.* Princeton, NJ: Princeton University Press.

Grèzes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. *Human Brain Mapping*, *12*(1), 1–19.

Guagnano, D., Rusconi, E., Umiltà, C. (2010). Sharing a task or sharing space? On the effect of a confederate in action coding. *Cognition*, *114*, 348–355.

Haken, H., Kelso, J.A.S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics, 51*(5)*, 347–356.*

Heed, T., Habets, B., Sebanz, N, & Knoblich, G. (2010). Others' actions reduce crossmodal integration in peripersonal space. *Current Biology, 20*, 1345–1349.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24*, 849–937.

Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition*, *27*(6), 949–961.

Hutchins, E. (1995). How a cockpit remembers its speeds. *Cognitive Science, 19(3)*, 265–288.

Huygens, C. (1673/1986). *The pendulum clock or geometrical demonstrations concerning the motion of pendula as applied to clocks.* Ames: Iowa State University Press (R.J. Blackwell, Trans.).

James, W. (1890). *The principles of psychology* (2 vols.). New York: Holt.

Kieras, D. & Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction*, *12*, 391–438.

Kladzky, R. L. (1975). *Human memory: Structures and processes.* San Francisco: W. H. Freeman & Co.

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: Theory and data. In B. Ross (Ed.), *The psychology of learning and motivation*, 54 (pp. 59–101), Burlington: Academic Press.

Knoblich, G., & Flach, R. (2001). Predicting the effects of actions: Interactions of perception and action. *Psychological Science, 12*, 467–472.

Kourtis, D., Knoblich, G., Wozniak, M., & Sebanz, N. (2014). Attention allocation and task representation during joint action planning. *Journal of Cognitive Neuroscience, 26*(10), 2275–2286.

Kourtis, D., Sebanz, N., & Knoblich, G. (2013). Predictive representation of other people's actions in joint action planning: An EEG study. *Social Neuroscience, 8*, 31–42.

Kourtis, D., Sebanz, N., & Knoblich, G. (2010). Favouritism in the motor system: Social interaction modulates action simulation. *Biology Letters, 6*, 758–761.

Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social connection through joint action and interpersonal coordination. *Topics in Cognitive Science*, *1*(2), 320–339.

Meyer, M., van der Wel, R.P.R.D., & Hunnius, S. (2013). Higher-order action planning for individual and joint object manipulations. *Experimental Brain Research*, *225*(4), 579–588.

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current Biology*, *20*(8), 750–756.

Neda, Z., Ravasz, E., Brechte, Y., Vicsek, T., & Barabasi, A. L. (2000). The sound of many hands clapping. *Nature*, *403*, 849–850.

Newell, A., Shaw, J. C., & Simon, H. A. (1958) Elements of a theory of human problem solving. *Psychological Review*, *23*, 342–343.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology, 9*, 129–154.

Ramenzoni, V.C., Davis, T., Riley, M.A., Shockley, K., & Baker, A.A. (2011). Joint action in a cooperative precision task: Nested processes of intrapersonal and interpersonal coordination. *Experimental Brain Research, 211*, 447–457.

Ramnani, N., & Miall, C. (2004). A system in the human brain for predicting the actions of others. *Nature Neuroscience*, 7(1), 85–90.

Ramsey, R., Hansen, P.C., Apperly, I.A., & Samson, D. (2013). Seeing it my way or your way: Frontoparietal brain areas sustain viewpoint-independent perspective selection processes. *Journal of Cognitive Neuroscience, 25*, 670–684.

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, *29*, 1046–1060.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination: Common ground and the coupling of eye movements during dialogue. *Psychological Science*, *18*(5), 407–413.

Richardson, M. J., Marsh, K. L. & Baron, R. M. (2007). Judging and actualizing intrapersonal and interpersonal affordances. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 845–859.

Richardson, M. J., Marsh, K. L., Isenhower, R., Goodman, J., & Schmidt, R. C. (2007). Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human Movement Science*, *26*, 867–891.

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews Neuroscience*, *11*, 264–274.

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: What other people see is calculated by low-level and early acting processes. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 1255–1266.

Schmidt, R.C., Carello, C., & Turvey, M.T. (1990). Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. *Journal of Experimental Psychology: Human Perception and Performance, 16(2)*, 227–247.

Schmidt, R. C., & O'Brien, B. (1997). Evaluating the dynamics of unintended interpersonal coordination. *Ecological Psychology*, *9,* 189–206.

Schmidt, R. C., & Richardson, M. J. (2008). Dynamics of interpersonal coordination. In A. Fuchs & V. Jirsa (Eds.), *Coordination: Neural, behavioral and social dynamics*. Berlin: Springer.

Schuch, S., & Tipper, S. P. (2007). On observing another person's actions: Influence of observed inhibition and errors. *Perception & Psychophysics, 69*, 828–837.

Searle, J. (1990). Collective intentions and actions. In P. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in communication*, Cambridge, MA: Bradford Books, MIT press.

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

Sebanz, N., Knoblich, G., & Prinz, W. (2003).Representing others' actions: Just like one's own? Cognition, 88, B11-B21.

Sebanz, N., Knoblich, G., & Prinz, W. (2005). How two share a task: Co-representing stimulus-response mappings. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 1234–1246.

Shockley, K., Santana, M. V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance, 29*(2), 326–332.

Simon, J.R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology, 81*, 174–176.

Spivey, M.J. (2007). *The continuity of mind*. New York, NY: Oxford University Press.

Tollefsen, D. (2005). Let's pretend! Children and joint actions. *Philosophy of the Social Sciences*, *35*, 75–97.

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.

van der Wel, R.P.R.D., & Knoblich, G. (2013). Cues to agency: Time can tell. In J. Metcalfe and H. Terrace (Eds.). *Agency and joint attention* (pp. 256–267). New York: Oxford University Press.

van der Wel, R.P.R.D., Sebanz, N., & Knoblich, G. (2014). Do people automatically track others' beliefs? Evidence from a continuous measure. *Cognition, 130*(1), 128–133.

van der Wel, R.P.R.D., Sebanz, N., & Knoblich, G. (2013). Action perception from a common coding perspective. In K. Johnson and M. Schiffrar (Eds.). *People watching: Social, perceptual and neurophysiological studies of body perception* (pp. 101–119). New York: Oxford University Press.

van der Wel, R.P.R.D., Sebanz, N., & Knoblich, G. (2012). The sense of agency during skill learning in individuals and dyads. *Consciousness and Cognition, 21*, 1267–1279.

van der Wel, R.P.R.D., Knoblich, G., & Sebanz, N. (2011). Let the force be with us: Dyads exploit haptic coupling for coordination. *Journal of Experimental Psychology: Human Perception and Performance, 37*(5), 1420–1431.

van Ulzen, N.R., Lamoth, C.J., Daffertshofer, A., Semin, G.R., & Beek, P.J. (2008). Characteristics of instructed and uninstructed interpersonal coordination while walking in pairs. *Neuroscience Letters, 432*, 88–93.

van Schie, H. T., Mars, R.B., Coles, M.G.H., & Bekkering, H. (2004). Modulation of activity in the medial frontal and motor cortices during error observation. *Nature Neuroscience, 7*(5), 549–554.

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

Vesper, C., van der Wel, R.P.R.D., Knoblich, G., & Sebanz, N. (2011). Making oneself predictable: Reduced temporal variability facilitates joint action coordination. *Experimental Brain Research, 211*, 517–530.

Vesper, C., van der Wel, R.P.R.D., Knoblich, G., & Sebanz, N. (2013). Are you ready to jump? Predictive mechanisms in interpersonal coordination. *Journal of Experimental Psychology: Human Perception and Performance, 39*(1), 48–61.

Wegner, D.M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G.R. Goethals (Eds.), *Theories of group behavior* (pp. 185–208). New York: Springer-Verlag.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*, 625–636.

Wiltermuth, S. S. & Heath, C. (2009). Synchrony and cooperation. *Psychological Science, 20*, 1–5.

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society, 358*, 593–602.