The Phenomenology of Joint Action: Self-Agency vs. Joint-Agency

Elisabeth Pacherie

The Phenomenology of Joint Action: Self-Agency vs. Joint-Agency

Elisabeth Pacherie

Pacherie, E. (2012). The Phenomenology of Joint Action: Self-Agency vs. Joint-Agency. In Axel Seemann (ed.), *Joint Attention*: New Developments, Cambridge MA: MIT Press, pp. 343-389.

14 The Phenomenology of Joint Action: Self-Agency versus Joint Agency

Elisabeth Pacherie

The military historian William H. McNeill opens his 1995 book *Keeping Together in Time: Dance and Drill in Human History* with a personal anecdote. In 1941, he was drafted into the U.S. army and sent to Texas for basic training. This involved a great deal of marching about and drilling on a dusty patch of the Texas plain. All concerned realized that this exercise was utterly useless given the facts of twentieth-century warfare, yet McNeill (1995) recalls the following:

> Marching aimlessly about on the drill field, swaggering in conformity with prescribed military postures, conscious only of keeping in step so as to make the next move correctly and in time somehow felt good.... A sense of pervasive well-being is what I recall; more specifically, a strange sense of personal enlargement; a sort of swelling out, becoming bigger than life, thanks to participation in collective ritual. (p. 2)

McNeill also points out the similarity of what he experienced to what happens in traditional communal dancing:

> "Boundary loss" is the individual and "feeling they are one" is the collective way of looking at the same thing: a blurring of self-awareness and the heightening of fellow-feeling with all who share in the dance. It matches my own recollection of what close-order drill felt like... (p. 8)

In these two examples of close-order drill and communal dancing, the joint action seems to bring participants a heightened sense of agency and a sense of we-ness at the expense of a well-defined sense of self. However, is this phenomenology characteristic of all joint actions? This is rather unlikely, given that joint actions come in a great variety of forms.

One can distinguish at least six relevant dimensions of variation in joint action. One concerns the number of participants involved in the joint action, from two at a minimum up to several million, as happened, for instance, in 2002 and 2003 with the huge street protests in many countries against the war in Iraq. A second dimension of variation concerns the more or less egalitarian versus hierarchical relationships among the participating agents. A third dimension of variation concerns the extent and form of the division of labor among coagents and thus the extent to which the roles they play are specialized rather than interchangeable. A fourth dimension of variation concern the nature of the interactions among participants, from purely virtual interactions, as in modern forms of telecommuting to work, to highly physical ones, as in communal dancing and close-order drill. A fifth relevant dimension concerns the transient versus long-term nature of the association formed by the participants. Two people who jointly help an old lady get up after she falls in the middle of the street may never have met before and may never meet again. In contrast, two acrobats who do a joint number may have trained together for years. Finally, some joint actions depend on complex institutions and involve activities heavily regulated by norms while others do not.

It is important to note also that while large-scale, long-term, hierarchical organization and institutional structure are features that tend to co-occur, their co-occurrence is not necessary, and these dimensions are at least partially independent. Thus, although the landing of the Allies in Normandy in 1944 was indeed a very large-scale joint action, involving over 156,000 troops, the participants of which acted as part of the epitome of hierarchical institutions, the military, and with a clear division of labor, other large-scale joint actions such as street protests don't have these features.

This chapter aims at investigating the phenomenology of joint action and at gaining a better understanding of (1) how the sense of agency one experiences when engaged in a joint action differs from the sense of agency one has for individual actions and (2) how the sense of agency one experiences when engaged in a joint action differs according to the type of joint action and to the role one plays in it. In recent years, there has been a surge of interest in the phenomenology of individual action, and there is now considerable evidence that the sense of agency we experience for an individual action relies on a multiplicity of cues related to different levels of action specification and control. If the same principle holds for joint actions—that is, if the same kinds of relations hold between mechanisms of action specification and control and mechanisms involved in the generation of the sense of agency—then, to get a better grip on the phenomenology of joint action, we need to know how the mechanisms of action specification and control involved in joint action differ from those involved in individual action. We also need to know how these mechanisms may differ according to the type of joint action under consideration. For that, in turn, we need to understand what specific requirements bear on joint actions as opposed to individual actions or to one type of joint action as opposed to another.

In the next section ("The Sense of Agency for Individual Actions: Sources and Mechanisms"), I offer a brief survey of recent, complementary models of how and where in the cognitive architecture the sense of agency is generated, pointing out the relations they draw between action specification and control mechanisms and processes involved in the generation of the sense of agency. Next, in the third section ("Small-Scale Joint Actions") and fourth section ("Beyond Small-Scale, Egalitarian Actions"), I discuss the specific requirements that bear on joint action—in particular, the requirements concerning the coordination of participants' actions with respect to their joint goal—and the cognitive mechanisms needed to ensure that these requirements are met. To keep things manageable, I only distinguish between two broad types of joint action: small-scale, egalitarian joint actions, discussed in the third section, and larger-scale, hierarchical joint actions, discussed in the fourth section. With the ground thus prepared, I plunge into the heart of the matter in the fifth section ("The Sense of Agency for Joint Actions"), where I discuss the factors influencing the strength or intensity of the sense of agency one experiences when engaged in joint action, the extent to which agency is experienced as joint agency, and whether it is at the expense of a sense of self-agency.

The Sense of Agency for Individual Actions: Sources and Mechanisms

Empirical research on (individual) agency has explored a number of potential cues to agency, and different cognitive models for agency have been proposed, ranging from high-level cognitive mechanisms to low-level sensorimotor mechanisms.

Some authors have tended to focus on high-level cognitive mechanisms, invoking a "central" interpretive system to explain our awareness of our own agency. According to this approach, the sense of agency is subserved by a holistic mechanism that is concerned with narrative self-understanding. Our sense of what, if anything, we are up to is based on the operations of a high-level integrative process that draws on the agent's self-conception and tries to put the best spin on things that it can. Such a conception has strong Dennettian overtones. We turn Dennett's intentional stance inward and treat ourselves as entities whose behavior needs to be made sense of in light of an implicit theory of ideal agency.

Many authors have expressed some sympathy with, and in some cases whole-hearted commitment to, the narrative approach. Interpreting split-brain studies in light of Dennettian (Dennett, 1992) themes concerning the role of narrative in self-interpretation, Roser and Gazzaniga (2004, 2006) have argued that the left hemisphere contains an interpreter, whose job it is to make sense of the agent's own behavior. The psychologist Louis Sass has suggested that schizophrenic patients with delusions of alien control no longer feel as though they are in control of their actions because "particular thoughts and actions may not make sense in relation to the whole" (Sass, 1992, p. 214), and Stephens and Graham (2000) have further developed his proposal. Peter Carruthers (2007) suggests that

> our awareness of our own will results from turning our mind-reading capacities upon themselves, and coming up with the best interpretation of the information that is available to it—where this information doesn't include those acts of deciding themselves, but only the causes and effects of those events. (p. 199)

Holistic themes also play an important role in Daniel Wegner's influential treatment of agentive self-awareness (Wegner, 2002, 2005). On the one hand, Wegner argues that the sense of agency is typically inferred from the existence of a match between a prior thought and an observed action, where the thought occurs just before the action, the thought is consistent with the action, and other potential causes of the actions are not present. On the other hand, he also notes that we perform many actions without the benefit of such previews, and he suggests, "Even when we didn't know what we were doing in advance, we may trust our theory that we consciously will our actions and so find ourselves forced to imagine or confabulate memories of 'prior' consistent thoughts" (Wegner, 2002, p. 146).

A wide array of evidence can be marshaled in support of this high-level account. When young children happen to achieve a goal by luck, they will say that they had intended the action that yielded that goal all along (Phillips, Baron-Cohen, & Rutter, 1998). Split-brain subjects are prone to confabulate accounts of actions that are generated by their right hemisphere (Gazzaniga & LeDoux, 1978). Data from subjects in altered states of consciousness also support the narrative approach. For example, bizarre behaviors performed in response to hypnotic suggestion are often accompanied by elaborate rationalizations and confabulations on the part of the agents (Moll, 1889). Finally, this approach derives support from a number of laboratory studies with normal subjects, in which it has been shown that the sense of agency can be modulated by priming and by various contextual parameters (Aarts, Custers, & Wegner, 2005; Wegner, Sparrow, & Winerman, 2004; Wegner & Wheatley, 1999).

In contrast to this high-level approach, a number of researchers have proposed that the monitoring of action execution is crucial for agency and that the sense of agency is generated by low-level mechanisms that exploit performance-related sensorimotor cues.

Tsakiris and colleagues have investigated the possibility that efferent signals sent to the motor system while implementing an intention provide such cues. In particular, they have proposed that efferent signals are used to generate accurate temporal and kinematic predictions about how and when particular body parts should move (Tsakiris & Haggard, 2005; Tsakiris, Haggard, Franck, Mainy, & Sirigu, 2005; Tsakiris, Prabhu, & Haggard, 2006). In support of that claim, they have demonstrated that self-recognition of one's own bodily movements crucially depends on efferent signals.

Another line of evidence for the role of efferent signals in generating a sense of agency involves "intentional binding," a phenomenon in which self-produced movements and their effects are perceived as being closer together in subjective time than they actually are (Haggard & Clark, 2003; Haggard, Clark, & Kalogeras, 2002). More specifically, when a voluntary act (e.g., a button press) causes an effect (e.g., a tone), the action is perceived by the agent as having occurred later than it did, and the effect is perceived as having occurred earlier. In contrast, when similar movements and auditory effects occur involuntarily rather than voluntarily, the binding effect is reversed and cause and effect are perceived as further apart in time than they actually are. The phenomenon of intentional binding suggest that the sense of agency is constructed at the time of the action itself, that it exploits efferent signals and is an immediate by-product of the motor control circuits that generate and control the physical movement.

Another mechanism appeals to internal forward models used for action control (Blakemore & Frith, 2003; Frith, Blakemore, & Wolpert, 2000a, 2000b). According to this proposal, forward models are fed an efference copy of actual motor commands and compute estimates of the sensory consequences of the ensuing movements. The predicted sensory consequences are compared with actual sensory feedback (reafferences). When there is a match between predicted and actual state, the comparator sends a signal to the effect that the sensory changes are self-generated, and when there is no match (or an insufficiently robust match), sensory changes are coded as externally caused. Indirect evidence for this model comes from studies demonstrating that discrepancies between predictions and sensory reafferences affect tactile sensations (Blakemore, Wolpert, & Frith, 1998; Blakemore, Wolpert, & Frith, 2000) and visual perception of one's own actions (Leube et al., 2003). Direct evidence is also provided by studies demonstrating that agency is gradually reduced as these discrepancies increase due to spatial deviations and temporal delays (Fourneret & Jeannerod, 1998; Knoblich & Kircher, 2004; Knoblich, Stottmeister, & Kircher, 2004; Leube et al., 2003; Sato & Yasuda, 2005; van den Bos & Jeannerod, 2002).

However, as several authors have pointed out (Gallagher, 2007; Knoblich & Repp, 2009; Pacherie, 2008), the results of some of these studies are open to alternative interpretations in terms of perceptual rather than sensorimotor cues. It is well-known that we have little awareness of the proprioceptive feedback associated with movements or even of the corrections we make during goal-directed movements (de Vignemont, Tsakiris, & Haggard, 2006; Fourneret & Jeannerod, 1998). Indeed, passive movements are associated with more activity in the secondary somatosensory cortex than active movements (Weiller et al., 1996). Frith (2005) even suggests that lack of proprioceptive experience may be one indicator that one is performing a voluntary act. The vast majority of our actions aim at producing effects in the environment, and we normally attend to the perceptual effects of our movements rather than to the movements themselves. It may therefore be that perceptual cues rather than sensorimotor cues are crucial to the sense of agency. Direct evidence for this view comes from an experiment of Fourneret and Jeannerod (1998) where subjects are instructed to move a stylus on a graphic tablet on a straight line to a visual target. Subjects cannot see their drawing hand, only its trajectory, visible as a line on a computer screen. However, the experimenter introduces a directional bias electronically so that the visible trajectory no longer corresponds to that of the hand. When the bias is small (less than 14 degrees), subjects make automatic adjustments of their hand movements to reach the target but remain unaware that they are making these corrections. It is with larger biases that subjects become aware of a discrepancy and begin to use conscious monitoring of their hand movement to correct for it and to reach the target. These results suggest that although discrepancies between predicted and actual sensory feedback are detected at some level since they are used to make appropriate corrections of the hand movement, they do not influence the sense of agency. Rather, subjects' sense of agency for the action seems to rely mostly on a comparison of the predicted and actual perceptual consequences of their action. As long as the trajectory seen on the screen matches sufficiently well the predicted trajectory, proprioceptive information is ignored.

Further evidence that perceptual cues may contribute more to the sense of agency than sensorimotor cues comes from pathologies (Jeannerod, 2009). For instance, patients with schizophrenia are impaired in explicitly judging whether they are in control of perceptual events but not impaired in automatically compensating for sensorimotor transformations between their movements and the resulting perceptual events (Fourneret et al., 2002). Frontal patients, like patients with schizophrenia, have a preserved automatic sensorimotor control, contrasting with impaired action awareness and conscious monitoring (Slachevsky et al., 2003).

All the models I briefly reviewed share a core idea. They appeal to a principle of congruence between anticipated outcome and actual outcome. Where they differ is on whether the cues used are primarily cognitive, perceptual, or sensorimotor. There is now, however, a growing consensus that these different models should be seen as complementary rather than as rival and that the sense of agency relies on a multiplicity of cues coming from different sources (Bayne & Pacherie, 2007; Gallagher, 2007; Knoblich & Repp, 2009; Pacherie, 2008; Sato, 2009; Synofzik, Vosgerau, & Newen, 2008). Thus, the conceptual framework I proposed (Pacherie, 2008), distinguishes between three hierarchically ordered intentional levels: (1) distal intentions, where the action to be performed (i.e., goals and means) is specified in cognitive terms, (2) proximal intentions, where it is specified in actional–perceptual terms, that is, in terms of the action schemas to be implemented and the perceptual events that will occur as a consequence, and (3) motor intentions, where it is specified in sensorimotor terms. As this model distinguishes between distal (D), proximal (P), and motor (M) intentions, I call it the DPM model, Comparisons of desired, predicted, and actual states at each of these three levels provide different cues to agency.

At present, these integrative frameworks still leave open a number of questions regarding the relative weight of different agency cues and the extent to which this weight can be modulated by the nature of the task, the attentional state of the agent, or the agent's level of expertise. To answer those questions, further empirical investigations are needed. However, these integrative frameworks all agree that the various cues exploited in generating the sense of agency for an action are signals and representations typically produced by action specification and control mechanisms and processes.

In what follows, I am assuming that the same kind of relationship holds for joint action, that is, that the sense of agency we experience for joint action is largely based on cues produced by the mechanisms of action specification and control at play in joint action. Thus, if we want to understand how the phenomenology of joint action differs from the phenomenology of individual action, we need to understand how the mechanisms of action specification and control involved in joint action differ from those involved in individual action. To understand that, in turn, we need to investigate what specific requirements bear on joint actions as opposed to individual actions or one type of joint action as opposed to another. To this task I now turn, starting with small-scale joint actions.

Small-Scale Joint Actions

In the broadest sense, the label "collective actions" can be used as a generic term to cover all cases where a certain global effect is the result of the actions of several individuals. Creating a traffic jam is a collective action in this sense, as a single motorist alone in the streets could not create a traffic jam. This is a collective action in the weakest possible sense since it is not even required that the participating agents have the goal to produce that outcome, that they coordinate to achieve it, or that they intend to act together. At the other end of the spectrum of collective actions are joint cooperative actions, where agents share the same goal, intend to act together, and coordinate their actions to achieve their shared goal.

Philosophers have tended to focus on the latter kind of collective actions, joint actions for short. Furthermore, their paradigmatic examples of joint actions tend to be small-scale, egalitarian joint actions, such as two people painting a house together, moving heavy furniture together, preparing a sauce together, or walking together. A number of prominent philosophers of action have proposed accounts aimed at capturing the features in virtue of which actions count as joint action (Gilbert, 1989, 1990, 2009; Tuomela & Miller, 1988; Tuomela, 2005; Searle, 1990, 1995; Bratman, 1992, 1993, 2009a, 2009b; Velleman, 1997). All agree that joint actions involve shared intentions (also sometimes called we-intentions, collective intentions, or joint intentions) and that a shared intention does not reduce to a mere summation of individual intentions, even supplemented by mutual beliefs or mutual knowledge. They disagree, however, on how best to analyze shared intentions. I will not enter into these debates here. Rather, my discussion will focus on Bratman's influential account (Bratman, 1992, 2009a, 2009b), reviewing its assets and pointing out some of its limitations.

Bratman on Shared Intentions

Bratman (1992) first identifies three features of joint actions, or, as he calls them in that paper, shared cooperative activities (SCA), that an analysis of shared intentions would have to account for:

1. *Mutual responsiveness* In SCA each participating agent attempts to be responsive to the intentions and actions of the other, knowing that the other is attempting to be similarly responsive. Each seeks to guide his behavior with an eye to the behavior of the other, knowing that the other seeks to do likewise.
2. *Commitment to the joint activity* In SCA the participants each have an appropriate commitment (though perhaps for different reasons) to the joint activity, and their mutual responsiveness is in the pursuit of this commitment.
3. *Commitment to mutual support* In SCA each agent is committed to supporting the efforts of the other to play her role in the joint activity.... These commitments to support each other put us in a position to perform the joint activity successfully even if we each need help in certain ways. (Bratman, 1992, p. 328)

None of these three features is by itself sufficient to make an activity an SCA, but, according to Bratman, taken together they are characteristic of SCAs. Bratman then argues that joint actions can be accounted for in terms of shared intentions.

However, how can shared intentions satisfy these requirements? With regard to the commitment to a joint activity, Bratman (1992, 2009a, 2009b) proposes that each of the participating agents should have an intention in favor of the joint activity, where to avoid circularity, the notion of a joint activity should be read in a cooperatively neutral way. Since Bratman construes commitment to a joint activity in a cooperatively neutral way, this commitment does not suffice to ensure that the activity that follows is an SCA. The originality of Bratman's analysis comes from the way in which he construes the two further features of mutual responsiveness and commitment to mutual support. Mutual responsiveness is analyzed in terms of interlocking intentions and meshing subplans. For an activity to be an SCA, it must be the case that the intentions of the participants interlock in the sense that each agent intends that the shared activity go in part by way of the relevant intentions of each of the other participants. Furthermore, each must also intend that this shared activity proceeds by way of subplans of the participants that mesh in the sense that they are corealizable.

Bratman analyzes the commitment to mutual support as the rational requirement that agents be disposed to help their partners play their role if their help is needed and not too costly to them. As a final condition on shared intentions, Bratman requires that there be common knowledge among the participating agents of all these conditions.

By conceiving of shared intentions as an interlocking web of intentions of individuals, Bratman moves away from the classical reductive analyses of collective action since he maintains that the crucial link among the attitudes of agents involved in joint action is not a purely cognitive link. Mutual belief or mutual knowledge does not suffice to ensure that intention is shared. What is crucial rather is the specific form of interdependence of the individual intentions of the participants.

Bratman's account is quite illuminating as an analysis of shared intentions for future joint activities of small, egalitarian, adult groups. Yet, it is unclear whether it can be generalized to other kinds of joint action. Some philosophers (Tollefsen, 2005; Butterfill, 2010) point out that Bratman's analysis presupposes that the participants have robust mind-reading and metarepresentational capacities and would not extend easily to the case of joint actions performed by young children or animals who lack those capacities. For lack of space, I won't discuss this issue here. Other philosophers (Kutz, 2000) have also expressed doubts that Bratman's analysis retains its plausibility when we turn to more complex cases of joint actions involving a high number of participants and/or embedded in institutional frameworks with structures of authority. I will consider this issue in the next section.

Even as an account of small-scale, egalitarian joint actions, Bratman's account can be seen as incomplete insofar as his focus is on shared intentions regarding future joint actions and, thus, on demands concerning the planning of joint actions rather than their execution. He provides illuminating analyses of the kind of attitudes and commitments participating agents must form in order to be said to share an intention to perform a certain joint action in the future, but he tells us very little about how joint actions are actually carried out and what capacities are required for their successful execution. Thus, while Bratman rightly insists that SCAs require mutual responsiveness not just of intention but also of action, he does very little to unpack what responsiveness in action amounts to and what capacities it involves. However, if we want get a grip on the phenomenology of joint action, we also need to understand how joint actions are able to unfold in time.

Until recently the cognitive and neural processes involved in joint action were little known. However, in recent years, major advances have been made and empirical data from both psychology and neuroscience have started to accumulate. Their exploitation may help philosophers extend their analyses beyond the level of distal intentions. In the remainder of this section, I will use the DPM model (Pacherie, 2008) as a guide for integrating these new empirical data and understanding how the characteristic features of SCAs find expression in the joint actions themselves rather than just in the distal intentions that (may) precede them.

In the same way that three different levels of intentions and control can be distinguished for individual actions, joint actions can be thought to involve a three-tiered hierarchy of intentions and control processes. We can call these intentions shared distal intentions (*SD-intentions*), shared proximal intentions (*SP-intentions*), and, for reasons I will explain shortly, not shared but coupled motor intentions (*CM-intentions*). I now examine the characteristics of the intentions involved at each of these three levels in turn, with particular attention to the cognitive abilities involved and the mechanisms thought to underlie them.

Shared Distal Intentions

Although joint actions can occur on the fly rather than being planned in advance and do not always involve SD-intentions, many do. Bratman's account is, I think, quite perspicuous as an account of SD-intentions for small-scale, egalitarian actions involving adult participants. Here, I rely on his analysis, simply pointing out the main commonalities and differences between distal intentions and control for individual action and for joint action. In the case of individual D-intentions, the agent (1) represents both the overall goal and the whole plan and (2) all he or she represents is to be performed by himself or herself. In contrast, in the case of joint actions, the participating agents (1′) represent the overall goal yet need not represent the whole plan but only their own subplans and the meshing parts of the subplans of others and (2′) some of what they represent is to be performed by others. Both (1) and (1′) are in need of some qualification. When I settle on a certain goal, I need not yet have a complete plan for achieving that goal, but I commit myself to form a plan that meets means–end consistency demands. Similarly, as Bratman points out, "I need neither know nor seek to know of all your subplans for us to have a shared intention; nor need we already have arrived at complete, meshing subplans" (Bratman, 1993, p. 121). Yet, we are committed to achieving our joint goal by way of subplans that mesh and thus are committed to coordinated planning.

Thus, the consistency constraints that bear on SD-intentions go beyond those on individual D-intentions. In both cases, the agents are expected to adjust their means to their ends and their plans to what they believe the world is like as well as to the wider framework of activities and projects in which they are also involved. In some cases, individual actions are directed at other people rather than at objects, and in planning their actions agents may need to take into account the intentions and actions of others. For example, the film *Enemy at the Gates*, where two snipers, a Russian and a German, play a game of cat-and-mouse during the Battle of Stalingrad, provides a vivid illustration of a sophisticated mutual adjustment of intentions and actions. Although this form of dyadic adjustment is also necessary for joint action, it is clearly not sufficient. Obviously, the two snipers in the film are not cooperating; theirs is a deadly competition. What is furthermore required in the case of joint action is that participating agents share a goal and understand the combined impact of their respective intentions on their joint goal and adjust them accordingly. The demand for triadic adjustment of plans thus constitutes a further consistency constraint specific to shared intentions and may be seen as their hallmark. Indeed, the minimal cooperative stability Bratman requires for shared intention is a distinctive echo on the control side of the triadic adjustment demand on the planning side.

Shared Proximal Intentions

The successful performance of joint actions requires not just that participating agents have a joint goal and meshing subplans that meet the consistency requirements on SD-intentions. It is also necessary that the agents be able to anchor these subplans to the situation of action and carry them out in a coordinated manner. To understand how this is possible, we need to acknowledge at least one further level of shared intentions, SP-intentions. So far, philosophers have had very little to say on what SP-intentions exactly involve and what cognitive capacities are needed to sustain them. Fortunately, psychologists and neuroscientists have recently started investigating the abilities needed for successful online coordination and the cognitive and neural processes underlying these abilities (Bekkering et al., 2009; Knoblich & Sebanz, 2008; Newman-Norlund et al., 2007a; Sebanz et al., 2006a; Sebanz & Knoblich, 2009).

I start with a brief characterization of what SP-intentions involve. For agents to share a proximal intention, the following should obtain: (1) agents each represent their own actions and their predicted consequences in the situation at hand (*self-predictions*), (2) agents each represent the actions, goals, motor and proximal intentions of their coagents and their consequences (*other-predictions*), (3) agents each represent how what they are doing affects what others are doing and vice-versa and adjust their actions accordingly (*dyadic adjustment*), (4) agents each have a representation (which may be only partial) of the hierarchy of situated goals and desired states culminating in the overall joint goal (*joint action plan*), (5) agents each predict the joint effects of their own and others' actions (*joint predictions*), and (6) agents each use joint predictions to monitor progress toward the joint goal and decide on their next moves, including moves that may involve helping others achieve their contributions to the joint goal (*triadic adjustment*).

Following Sebanz et al. (2006a), we may distinguish three main types of abilities on which SP-intentions depend for their formation and operation. First, to anchor a SD-intention into the situation of action in order to generate corresponding SP-intentions, agents need to be able to form shared perceptual representations of the situation of action. Second, they need to be able to corepresent the actions and proximal intentions of other agents as well as their own to make possible dyadic online adjustments. Third, they need to be able to integrate the predicted effects of their own and others' actions in relation to the common goal to ensure the possibility of triadic online adjustments. Let us examine several mechanisms that have been proposed to underlie these abilities.

Several researchers have suggested that joint attention provides a basic mechanism for sharing representations of objects and events and thus for creating a "perceptual common ground" in joint action (Tomasello, 1995, 1999; Tomasello & Carpenter, 2007; Tollefsen, 2005; Sebanz et al., 2006a). The phenomenon of joint attention involves more than just two people attending to the same object or event. At least two additional conditions must obtain. First, there must be some causal connection between the two subjects' acts of attending (causal coordination). Second, each subject must be aware, in some sense, of the object as an object that is present to both; in other words the fact that both are attending to the same object or event should be open or mutually manifest (mutual manifestness). Empirical evidence indicates that although causal coordination and an understanding of what others are seeing are abilities found in several primate species, mutual manifestness and thus actual attention sharing may be unique to humans (Tomasello & Carpenter, 2007).

Joint attention plays two important roles in SP-intentions. First, the joint action plan must be anchored into the situation of action. For that, it is necessary that the objects to be acted upon, their location as well as the location of possible obstacles, be identified by the coagents and thus that they track the same objects and features of the situation and be mutually aware that they do. Second, once the joint action unfolds, coagents must coordinate their respective actions and, for that, must pay attention to what others are doing or about to do. Knowing what others are attending to in a particular situation provides important cues about their subsequent actions. Joint attention would thus play a crucial role in ensuring that the meshing of subplans translates into a corresponding meshing of actions.

However, for joint actions, elementary or not, to be possible, joint attention is not enough. It is also necessary that agents be able to corepresent the actions, goals, and proximal intentions of other agents as well as their own. A number of recent theories—the common coding theory (Prinz, 1997), the motor simulation theory (Jeannerod, 1997, 2006), and the motor resonance theory (Rizzolatti & Craighero, 2004)—converge on the idea that action observation can support the understanding of goals and intentions. These theories postulate an interface between perception and action such that the perception of an action leads to the activation of a corresponding action representation in the observer's action system. These theoretical insights are supported by a wealth of empirical findings. Using single-cell recording techniques, Rizzolatti and his coworkers discovered that a subpopulation of neurons in the ventral premotor area F5 of macaque monkeys is activated both when a monkey executes certain goal-directed hand or mouth movements and when it sees similar goal-directed movements performed by conspecifics or by human experimenters (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). The perceptual properties of mirror neurons (MNs, for short) appear to "mirror" their motor properties—hence, their name. Brain imaging studies provide evidence for the existence of a corresponding "mirror system" in humans, a set of brain regions activated both when an agent performs an action and when he or she observes actions of the same class performed by others (Decety & Grezes, 1999, 2006). The existence of such a mirror system in humans is also supported by behavioral experiments on motor interference, where observation of a movement is shown to degrade the performance of a concurrently executed incongruent movement (Brass, Bekkering, Wohlschlager, & Prinz, 2000; Kilner, Paulignan, & Blakemore, 2003).

Investigations of mirroring systems in humans have yielded evidence that their activity is involved in the execution and observation of a wider class of actions than in nonhuman primates, including intransitive actions (Buccino et al., 2001). Brain imaging results also show that mirror regions in humans may be associated with imitation and language (Carr et al., 2003; Fadiga et al., 2002; Iacoboni et al., 1999; Skipper et al., 2005). These findings suggest that mirror systems in humans involve more than just the kind of circuitry associated with MNs in monkeys. Rather, circuitry homologous to that of the macaque appears to be embedded in more extended systems within the human brain (Oztop et al., 2006).

Many have claimed that MNs and mirror systems support action understanding. However, one can mean rather different things by “action understanding,” and it is unlikely that macaque-like MN circuitry and the more complex mirror systems in humans support the same forms of action understanding. On a modest reading of “action understanding,” motor resonance would support action understanding insofar as it would allow the observer to retrieve the underlying goal of the observed action. But here the notion of goal should be understood in a correspondingly modest way, that is, as the immediate motor goal of the action. As pointed out by Sebanz et al. (2006a), this could help to establish procedural common ground in joint action. By sharing representations of actions and their motor goals, agents would be in a position to understand what their coagents are currently doing.

However, as Sebanz et al. (2006a) also remark, to interact successfully with others, knowing what they are currently doing may not be sufficient; it may also be crucial to be able to predict the outcomes of others’ actions and what they are going to do next. There is also evidence that motor resonance supports outcome prediction (Wilson & Knoblich, 2005). Indeed, as Csibra and Gergely (2007) point out, one way in which it does is rather trivial. If motor resonance supports goal attribution and a goal represents a state or an event subsequent to the action it belongs to, then goal attribution to not yet completed actions implies, by definition, a specific prediction. This type of prediction is what they call an “action-to-goal” prediction. It could be critical for joint actions in which goals are constantly in flux and where success requires that coactors perform complementary actions in quick succession. The fact that a fair proportion of MNs are broadly or logically congruent (Fogassi & Gallese, 2002), responding to observed actions similar or causally related rather than identical to the performed actions they also code for, suggests that they are relevant for complementary action. Indeed, a recent brain imaging study (Newman-Norlund et al., 2007b) found that the human MN system was more active during complementary compared to imitative actions.

A second type of prediction, less trivial and perhaps even more relevant to SP-intentions, involves action anticipation. These goal-to-action predictions, as Csibra and Gergely (2007) call them, would allow one to anticipate the observed actor’s next actions. Here, however, the notion of a goal cannot be interpreted in the same modest way as in action-to-goal predictions, that is, as simply an elementary motor goal. More complex inferential work is needed. The observer must take into account the motor goal of the perceived action together with contextual factors (the situation in which this motor action takes place) in order to first infer the superordinate goal toward the satisfaction of which the currently perceived action may contribute and then infer what further actions are needed to achieve that superordinate goal. Thus, although basic motor resonance and indeed individual MNs may support action-to-goal predictions, goal-to-action prediction requires the involvement of more complex processes of teleological reasoning. According to Csibra and Gergely (Csibra & Gergely, 1998; Gergely & Csibra, 2003), teleological reasoning is based on the assumption that agents engage in the most efficient course of action to achieve their goal within the situational constraints given—what they call the principle of rational action. Thus, when observing an ongoing action, teleological reasoning can be used to infer the likely goal of the action by assessing what end state would be efficiently brought about by the action given the particular situational constraints, or, if the goal is known, to generate an action prediction by inferring what the most efficient course of action toward the goal state would be in the given situation. As Csibra and Gergely point out, teleological reasoning is a very flexible tool in action understanding. However, it will lead to legitimate conclusions only if (1) the observed actor's behavior approximates the ideal of efficiency and (2) the observer is able to recruit relevant background knowledge about the physical constraints of the situation and of the actor. Since biological systems tend to conserve energy, condition (1) is likely to hold, and thus teleological reasoning is likely to be a computationally viable way of teleological action understanding (Baker et al., 2006). However, predictions may still go wrong if the observer has insufficient knowledge about the constraints of the actor or the situation and thus fails to meet condition (2).

If we come back to the insight that we understand an observed action by activating a corresponding action representation in the observer's action system and if we accept the idea that human action systems are organized hierarchically into three main levels of representation and control, the view at which we arrive is the following. The more basic form of action understanding involves representing the motor intentions underlying observed actions; this understanding recruits the same mechanisms and processes that are involved in the formation and control of the observer's own motor intentions, in particular the forward and inverse models that operate at the motor level. In contrast, the more demanding form of action understanding we have discussed involves inferring and representing the proximal intentions of the observed agent; to do so, it recruits the mechanisms and processes involved at the level of proximal intentions.

Recall that at the level of proximal intentions, action and goal representations are more abstract than at the motor level yet are still firmly anchored on a particular situation of actions rather than being detached. To understand the proximal intentions of others, it is thus not enough that one be able to retrieve the immediate motor goal of an observed movement; one must also infer its possible significance given further information about the situation in which it occurs. Moreover, given that even in a single situation there may be a number of different action sequences an elementary motor act may be part of, to further narrow down the range of possibilities, it may be crucial to identify the features of the situation the actor is attending to. Thus, to form reliable representations of coagent proximal intentions based on action observation would involve not just recruiting for simulation purposes the forward and inverse models that operate at the P-level but also filtering input to those simulation processes using one's knowledge of what the coagent is attending to.

Sebanz et al. (2006a) describe another important means to predict others' actions and intentions: task sharing. By knowing what another's task is—that is, knowing the stimulus–response contingencies of that task—one can predict what he or she is likely to do. Empirical evidence shows that when subjects know these stimulus–response mappings, they generate a representation of the appropriate action following stimulus presentation but in advance of action observation (Kilner et al., 2004; van Schie et al., 2004). Furthermore, a series of recent studies (Sebanz et al., 2005; Sebanz et al., 2006b, 2007) showed that actors form shared representations of tasks quasi-automatically, even when it is more effective to ignore one another. Shared representations of tasks as well as shared representations of proximal intentions (rather than simply motor intentions) thus allow coagents to extend the temporal horizon of their own planning, by making it possible for them to anticipate others' future actions and prepare responses to these future actions.

For shared representations of actions and tasks to foster coordination rather than create confusion, it is important that agents also be able to keep apart representations of their own and of others' actions and intentions. Unless it is clear who is doing (or preparing to do) what, coagents cannot efficiently plan their next moves. Although the exact mechanisms through which self–other distinction is achieved are not yet well understood, there is growing brain imaging and clinical evidence that the right parietal cortex and the insula are strongly implicated in this process of self–other distinction (Ruby & Decety, 2001; Farrer & Frith, 2002; Farrer et al., 2003; Jeannerod & Pacherie, 2004). In particular, existing data indicate that activation in the right inferior parietal lobule is negatively correlated, and activation of the insula positively correlated, with self-agency. Since both areas are involved in various forms of mapping and integration of multimodal information, agency attribution and self–other distinction appear to be based on processes of comparison of information from different sources, including interoceptive, exteroceptive, and motor feedback signals.

I have so far discussed two kinds of abilities successful action depends on: the ability to share perceptual representations of the situation of action and the ability corepresent the actions and proximal intentions of coagents while maintaining a self–other distinction. A third kind of ability is also required, which is perhaps the most crucial—namely, the ability to integrate the predicted effects of one's own and others' actions in relation to the joint goal. Joint attention and corepresentations of others' actions and intentions can support both competitive and cooperative interactions, but this third kind of ability is where the difference between cooperation and competition lies. Unfortunately, however, this ability is also the least well understood. It is, as Knoblich and Sebanz (2008) put it, "critical and miraculous at the same time"(p.2025). Some recent neuroimaging studies (Newman-Norlund et al., 2007a, 2008) raise the possibility that right inferior frontal activations are related to integration processes; however, other interpretations of these activations in terms of inhibition processes are possible (Brass et al., 2001; Brass et al., 2005). Since empirical data are still scarce, the suggestions I have to offer regarding integration are perforce highly speculative.

Bratman (2009a) spells out that requirement at the level of distal intentions in terms of a web of intentions allowing an agent to relate his or her own intentions and the intentions of his or her coagents to an intention in favor of a joint activity. However, how does that translate at the level of proximal intentions, and what form does that intentional structure take? For there to be a joint action, coactors have to be able to relate and adjust their own actions and the actions of their partners not just to one another but to the joint action. This requires that agents be capable of explicitly representing the instrumental relation of their and their coagents' individual actions to the joint action, and this, in turn, requires that agents form a detailed representation of their joint goal that carves it, so to speak, at its instrumental joints. I therefore propose that the representation of the joint goal that agents form at the level of SP-intentions consists in a representation of a hierarchy of situated goals. This representation would be more specific than the kinds of plans that would be attached to SD-intentions insofar as goals are indexed to a specific situation and goals and subgoals can be represented more concretely as desired states in that situation. Suppose, for instance, our joint goal, as it could be represented at the level of SD-intentions, is to rearrange the furniture in the living room by inverting the position of the dining space and of the television corner. At the level of proximal shared intentions, this goal can be specified more concretely as moving this table from here to there, placing the sofa along the wall facing the window, and so forth, and from this situated representation a hierarchy of subgoals can be derived such as first clearing obstacles off the way, unplugging the TV set, and so on. Note that for this representation of a hierarchy of situated goals to be shared, coagents should jointly attend to the situation. At the same time, this representation remains more abstract than representations agents may form of their own actions and of those of their coactors since by itself it neither specifies the precise means to be employed to achieve the various situated goals and subgoals nor who is to do what.

To relate and adjust their own actions and the actions of their partners not just to one another but to the joint action, agents should be capable of explicitly representing the instrumental relation of their individual actions to the situated joint goal structure. This leads to increased demands on executive control. Actors need to do more than just keep track of who's doing what and of how what others are doing affects what they themselves are doing or are going to do. They must also keep track of how what each is doing contributes (or, if their actions are unsuccessful, fails to contribute) to the achievement of goals and subgoals within the joint goal hierarchy, thus monitoring progress toward the achievement of the overarching joint goal and allowing them to plan their next moves, including moves that involve helping others achieve their contributions to the joint goal.

When roles have not been distributed in advance, it is important to figure out what others can do in order to decide whether or not to take care oneself of a given subgoal, let others take charge, or do it together. This may involve using prior information one already has about one's own and one's coagents' respective skills. In some cases, one may also have to exploit online perceptual information. In the example given earlier of people rearranging furniture in a room, this would involve taking into account the bulk and weight of the various pieces of furniture, as well as the location of the coactors relative to various objects: it takes two to move a heavy sofa, but one can move a chair on one's own, and if we're moving chairs it makes sense for me to take care of the chairs near where I stand and for you to take care of the chairs near you. Interestingly, in a series of experiments, Richardson and colleagues (Isenhower et al., 2005; Marsh et al., 2006; Richardson et al., 2007a) have shown that when acting together, people also take into account the motor affordances of their coactors. In these experiments, they paired subjects with different arm spans and asked them to lift planks off a conveyor belt. The planks could only be touched at their extremities and varied in length such that some could be lifted by a single individual and others only by two individuals. The planks were presented in ascending, descending, or random order of length. The transition between one-person lifting and two-person lifting during the ascending and random order tended to occur around the time the smaller participant could no longer comfortably lift the planks alone. The greater the difference in arm span between the two participants, the earlier the transition, suggesting that in deciding what to do the participants with the longer arms were taking into account the motor affordances of their partner.

Coupled Motor Intentions

Finally, for some joint actions to proceed successfully, it is not enough that coactors share a representation of a hierarchy of situated goals and converge on the distribution of roles; it is also necessary that their actions be very precisely coordinated in time and space. Ballroom dancing or rowing are fitting examples. In such cases, we need to appeal not just to SP-intentions but also to CM-intentions.

There is evidence that basic and unconscious bodily entrainment mechanisms may help achieve synchronization. Thus, two people sitting next to each other in rocking chairs will unconsciously synchronize their rocking frequency and do so even when they have chairs with different eigenfrequencies (Richardson et al., 2007b). Similarly, people interacting tend to nonconsciously mimic each other's gestures, postures, and mannerisms, and this unconscious mimicry has been shown to enhance the smoothness of interactions and foster liking (Chartrand & Bargh, 1999). Such entrainment mechanisms may thus facilitate the formation of CM-intentions. I have already discussed one way in which motor intentions may become effectively coupled: basic motor resonance mechanisms involving broadly congruent mirroring allow for "action-to-(motor)-goal" prediction and for the automatic activation of complementary actions by the observer. More generally, the extent to which motor coupling can be achieved appears to depend on the degree of similarity between the motor repertoires of the agents as well as on their level of expertise. Thus, a study by Keller et al. (2007) found that pianists duet better when they play with themselves, that is, are better at synchronizing with recordings of their own past performances than with others' recordings. This finding seems to indicate that tight action coordination and synchronization requires that coagents have similar internal models not just at the proximal level but also at the motor level.

Note that I speak here of *coupled* rather than shared motor intentions. Here's why. I insisted earlier that one essential feature of both shared distal and proximal intentions is that coagents having these intentions have a representation of a joint goal as such in addition to representations of their individual intentions and actions and have mechanisms of triadic adjustment with respect to the represented joint goal. I don't think existing empirical evidence allows us to posit the existence of motor representations of joint goals or of mechanisms of triadic motor adjustment. Rather, what we have at this level are simpler mechanisms of dyadic adjustment. When the motor intentions of the coagents are embedded within SP-intentions and placed under their control, thus ensuring that they attend to the same aspects of the environment and elicit parallel motor simulations, motor intentions can become coupled so as to promote the joint proximal goal. CM-intentions whose coupling is modulated by an SP-intention would thus mimic shared motor intentions.

In addition, recent studies (Wiltermuth & Heath, 2009) indicate that motor synchronization can foster cooperation within groups by strengthening group cohesion. Thus, in one of these studies, an experimenter led thirty participants in groups of three on walks around campus. In the synchronous condition, participants walked in step. In the control condition, they walked normally. After their walk, participants completed a questionnaire designed to convince them that they had finished the experiment. In an ostensibly separate experiment, a second experimenter had them play an economic game, the Weak Link Coordination Exercise, where different amounts of cooperation and free riding are possible. Participants who had walked in step cooperated more than those who had not walked in step. Participants in the synchronous condition also indicated stronger feelings of connection with and trust in their counterparts than did those in the asynchronous condition.

To recap, I have argued in this section that to understand how joint actions are able to unfold over time and be successful completed, it is not enough to postulate Bratman-like SD-intentions. To account for the successful performance of small-scale joint actions, we need to acknowledge at least one further level of shared intentions, SP- intentions, responsible for online dyadic and triadic adjustments. To account for joint actions whose successful performance requires very precise spatiotemporal coordination, we also need to acknowledge CM-intentions. SP-intentions and CM-intentions are supported by a number of cognitive mechanisms and processes, including, for the former, joint attention, motor resonance, and teleological reasoning, and, for the latter, bodily entrainment mechanisms.

Beyond Small-Scale, Egalitarian Actions

In the previous section, I concentrated on small-scale, egalitarian joint actions involving face-to-face interactions. My focus was on what was needed, beyond simply SD-intentions, to ensure their successful performance. In face-to-face interactions, the coordination of individual actions in the pursuit of a joint goal is made possible in large part by various forms of exploitation of perceptual information. Thus, joint attention mechanisms use perceptual information to determine what is and what is not common perceptual ground among coactors while motor resonance mechanisms and teleological reasoning use perceptual information about the actions others are performing and about situational constraints to infer their goals and future actions and predict their consequences. Yet, if these were the only cognitive tools we had at our disposal to promote coordination, there would be sharp limitations to the kind of joint actions we could successfully engage in.

First, since these cognitive tools exploit perceptual information, they can be of no help unless a certain amount of common perceptual information is indeed available to coactors. Second, even when common perceptual information is available, there are limits to our processing capacities. An agent may be able to simultaneously track what a small number of other agents are currently attending to, but when the number of agents and the number of different things they are attending to increase, this capacity soon finds its limits. Our capacity to co-represent the actions, goals, and proximal intentions of other agents we observe acting encounters similar limitations. Moreover, understanding of actions through motor resonance or mirroring works only to the extent that the observed actions are part of the action repertoire of the observer.

Yet, human agents have been able to overcome these limitations. They engage in joint actions involving large numbers of coactors. They engage in joint actions where they don't have common perceptual grounds and where interactions are virtual rather than physical. They engage in joint actions where they play specialized roles that are not interchangeable. The questions I am concerned with in this section are the following. How is coordination toward a joint goal achieved in such cases? What cognitive capacities does it tap? How should we revise a Bratmanian account of joint actions to accommodate these kinds of joint actions?

I start with an exploration of the different kinds of "coordination tools" that are involved in these more complex forms of joint action. To do this, I discuss at some length the fascinating example of the symphonic orchestra. An orchestra playing a symphony is certainly a prime instance of joint action. This kind of joint performance has features in common with the simpler forms of joint actions we have been concerned with so far: the members of the orchestra are in the same physical location, they share a fair deal of perceptual ground, and the success of their collective performance depends in part on a very tight temporal coordination of their respective individual actions. Yet, there are also important disanalogies.

Chief among the features that distinguish the performance of a symphonic work from the performance of simpler joint actions are the numbers of agents involved and the complexity of the task they set for themselves. A modern symphony orchestra has around eighty to one hundred musicians, playing from ten to over twenty different instruments. A symphonic work is a highly complex musical piece where different instruments or groups of instruments play different but simultaneous musical "lines" whose tempo and dynamics evolve in time. How can the musicians possibly succeed? What are the coordination tools that make it possible for them to hold their performance together? The score, orchestra hierarchy, and conductor, as well as ensemble practice and rehearsals, are instances of the different classes of coordination tools that make complex joint actions possible. Let us now examine the functions of these tools and see what they contribute to the various levels of intentions and control involved in joint action.

The musical *score* provides an explicit representation, both material and public, of the joint goal and action plan, assigning to each agent his or her part in the joint action and providing for each a detailed script of what to do. If people had to devise plans from scratch every time they are about to engage in joint action, these would be cognitively very costly and time-consuming, and we most probably would not witness that many performances of joint actions. Instead, very often, people rely on preestablished scripts, where plans and subplans that mesh are delineated and provide effective blueprints for the coactors. The musical scores used in Western classical music are perhaps an extreme case (in part because performing the music written on the score is the goal of the joint activity, not a means toward some further goal, and in part also because this musical tradition puts special emphasis on fidelity to the score and to the intentions of the composer), but examples also abound outside the domain of music or of the performing arts. Surgical teams in operation theaters have well-established procedures where each member of the team knows exactly what he or she has to do; the same is true of fire fighters, sailors, and in general of most activities where people work as a team. Preestablished scripts thus function as ready-made SD-intentions, reducing demands for the negotiation of plans and subplans among coagents.

However, even the most precise scores or scripts cannot anticipate all the particulars of the situations where they will be put to use and are not so specific that they leave no room for interpretation. But then again, it would be cognitively costly and time-consuming if all the participants needed, so to speak, to sit around a table and start deliberating as to how the script is to be interpreted and adapted to the situation at hand. Having a *hierarchical organization* is a way of curtailing this process. In the orchestra, the principal of each section is responsible for making decisions concerning his or her section. For instance, if needed, the principal cellist decides on the bowing movements for all the cellists so that, when they are playing tutti, they all bow up and down together. Decisions that concern the whole orchestra are taken by the conductor. The conductor's job is to interpret the intentions of the composer, which means choosing general levels of tempo and volume as well as supervising all the fine shading. Thus, the transition from (possibly ready-made) SD-intentions to SP-intentions is not entirely up to the individual participants but requires decisions to be made at various levels of the orchestra hierarchy.

However, even with all these matters settled and decisions made, precise coordination of the whole ensemble is not achieved instantaneously. Rather, it is a skill that needs to be honed, and that may only be achieved through *ensemble practice and rehearsals*. Within sections, coordination is facilitated by the fact that apart from the principal who may be called to play solo parts, musicians usually all play the same part on the same instrument. Indeed, the fact that, within each string section, players use the same bowing movements is not just a matter of visual aesthetics. Players in the same section are seated together in the orchestra, within sight of one another, thus allowing bodily entrainment mechanisms to help achieve synchronization.

When the different instrumental sections rehearse together, the role of the *conductor* becomes crucial. Musicians typically have visual access to only a fraction of their fellow players and, depending on their position in the orchestra, receive auditory feedback from the joint performance that is partial and variously distorted. Think, for instance, of the players seated next to the cymbalist! The conductor, in contrast, stands on a podium facing the orchestra. He has visual and auditory access to all the musicians and all the musicians see him or her. The conductor's role is that of a central coordinator. Leaving aside the finer stylistic aspects of the performance, the conductor's most basic responsibilities involve rhythmic and musical coordination as well as the quality and balance of sound. The conductor's job is to ensure that all the members of an orchestra start together and stay together, that individual players or sections make their musical entrance at the right moment, and that different but simultaneous musical "lines" are at the proper volume levels relative to their importance and one instrument or group of instruments doesn't inadvertently drown out any others.

The more complex a joint action, the more information is needed to keep it on track. This gives rise to the twin problems of information gathering and information processing. How can it be ensured that the coagents get access to the information they need to keep their actions on course? How can it be ensured that their information-processing load remains manageable? The conductor—as well as, to some extent, the hierarchical organization of the orchestra—provides an innovative solution to this problem, a solution that relies on a new kind of division of labor, new ways of communicating information, and new forms of commitments.

First, although, in simple joint action, there can be a division of labor in the sense that coagents may be assigned different but complementary tasks, they nevertheless all remain equally responsible for mutual responsiveness and support with respect to the joint goal and for what I called dyadic and triadic adjustments in the previous section. In contrast, the division of labor that creates a role for the conductor involves a redistribution of monitoring and control tasks. Musicians playing a symphony together may retain responsibility for local aspects of coordination, but the conductor is in charge of global coordination through monitoring and controlling all the individual contributions to the joint performance. Instead of each doing their part while monitoring what all the others are doing and controlling their actions accordingly, the players lighten their cognitive load by delegating the monitoring to the conductor and taking his or her cues in order to control their actions.

Second, the information conveyed to the musicians by the conductor in order to help them achieve coordination is not the kind of brute natural information (perceptual feedback, observation of others' actions) that agents use to coordinate their actions with those of others in simpler situations. The role of the conductor does not reduce to simply relaying the information that the conductor's privileged position on the podium gives him or her access to. Saying that the conductor is a central monitor and controller means that it is the conductor's responsibility to evaluate the information he or she receives, to compare it to the intended joint effect as determined by the score and his or her interpretation of it, and to give instructions to the musicians to adjust their performance accordingly. To convey those instructions, the conductor relies on a system of communicative signals. At least during concerts, these signals cannot be verbal, for they would interfere with the musical performance, but the communication system used by the conductor is nevertheless in part conventional. For instance, the primary function of the right hand holding the baton is to beat time, with the downbeat of the hand indicating the first beat of the bar, and changes in dynamics can be indicated in a variety of ways, such as changes in the size of conducting movements, upward or downward motion, or leaning toward or away from the performers.

In a nutshell, the various coordination tools we just examined help reduce the otherwise impossible demands on the cognitive resources of agents involved in complex joint action. They do so at each of the three levels of action representation and control. At the distal level, preestablished scripts, such as scores, help dispense with long negotiations and adjustments of plans and subplans. At the next level, where decisions have to be made as to how best to translate distal intentions into proximal ones, the existence of a hierarchical organization helps simplify the process of decision making, with leaders at various levels of the hierarchy given responsibility for decisions. Similarly, at the motor level, various coordination tools can be used to ensure a proper coupling of the motor behavior of agents. The orchestra conductor stands out as a particularly interesting object of study, both because the conductor plays important roles at all three levels of action representation and control and because he or she epitomizes the kind of division of labor emblematic of complex joint actions. All complex joint actions involve a degree of separation between executive tasks and control and monitoring tasks, not equally distributed among participating agents. In the case of the conductor, the separation is complete: the conductor doesn't contribute a sound to the musical performance but is responsible for monitoring and controlling the performance of the whole orchestra.

Finally, it should be noted that the hierarchical organization and new form of division of labor typical of complex joint actions transform the nature of the commitments taken by the coactors. If two people are painting a house together, it is reasonable to attribute to each, as proposed in Bratman's account, the intention "that we paint the house together," for their intentions are highly interdependent and both of them are responsible for the planning and meshing of subplans that will ensure the desired joint outcome. However, Kutz (2000) points out the following:

> It would ring false to attribute to an individual cellist in an orchestra the intention that "we play the Eroica" … Rather, it is far more natural to attribute to the cellist an intention to perform his or her part in the symphony.… In contrast, we might say of a conductor … that [he or she] intends that his or her group perform [the Eroica] given his or her ability to influence this total outcome. (p. 23)

In small-scale, egalitarian joint actions, agents are equal contributors to the shared intention and joint action. They are all equally involved in the choice of the main goal of their joint action, and they are all equally responsible for the planning toward that goal and the meshing of subplans. The dependence relations between their intentions are symmetrical. In particular, according to Bratman's analysis (Bratman, 2009), agents each intend that we *J* in part because we believe the other so intends; agents each believe that their successfully *J*-ing depends on the persistence of both their intentions, and they each believe that the persistence of their own intention depends on their continued knowledge of the persistence of the other's intention. The division of labor and hierarchical organization typical of complex joint action lead to differential contributions to joint activity, some marginal and others crucial, and to asymmetrical dependence relations among agents' intentions. For instance, musicians in the orchestra have very little influence on the choice of the musical works the orchestra is to perform, nor are they responsible for the choice of interpretation, or in charge of planning toward securing the group outcome or organizing rehearsals. All of those tasks are the responsibility of the conductor, whose planning and action is aimed at the goal that the orchestra together perform the musical work he or she has chosen and interpret it in the way he or she has decided. These differences suggest that we should attribute different types of intentions to the conductor (or more generally agents high in the hierarchy) and the orchestra musicians (or agents at lower levels of the hierarchy).

According to Kutz (2000), although the conductor can be attributed an intention with respect to producing the total outcome or activity (that we play the Eroica) given his crucial role in planning and acting toward this total outcome, it would be unwarranted to attribute an intention whose scope includes the entire performance to the orchestra musicians whose contributions are more marginal. Rather, their planning and action are directed toward the goal of performing their roles, and thus we only need attribute to them a participatory intention, that is, an intention to do their part in the collective act. Kutz's account entails that the dependence relations among the intentions of large hierarchical groups will be asymmetrical. The participatory intentions of the orchestra musicians are subsidiary to the intentions of the conductor: their intending to practice their instrument's part in the Eroica, rather than, say, the Pastoral, to play fast or slow, or to attend a rehearsal on Friday at 10 a.m. are causally dependent on the conductor's intentions and plans, but the converse does not hold. Their respective commitments will also differ. In small-scale egalitarian actions, all agents are committed to the joint activity. In contrast, in larger hierarchical groups, the conductor (or, more generally, agents high in the hierarchy) is committed to the joint activity and thus to planning toward the whole outcome while the musicians (or, more generally, participants lower down the hierarchy) are committed to doing their part in the collective activity. In small-scale egalitarian actions, all agents share the same commitments to mutual responsiveness and mutual support. In other words, solving coordination problems is their shared responsibility. In larger hierarchical groups, someone, for example, the conductor, has the responsibility to resolve coordination problems. For participants lower down the hierarchy, the commitments to mutual responsiveness and mutual support give way to a commitment to responsiveness to the leader's indications.

The Sense of Agency for Joint Actions

The discussion that follows is premised on the idea that, as is the case with individual actions, the sense of agency we experience for joint action relies on a multiplicity of cues related to different levels of action specification and control. However, the mechanisms of action specification and control involved in joint action are typically more complex than those present in individual actions. Thus, to understand how the phenomenology of joint action might differ from the phenomenology of individual actions, we need to take into account the specific requirements that bear on joint actions and the constraints these requirements impose on action specification and control processes. In the last two sections, I discussed these requirements, in particular the requirements for dyadic and triadic adjustments of intentions and actions among agents, as well as a range of cognitive tools we use to try and meet them.

In an investigation of the phenomenology of joint action, we should consider the issue of what factors influence the strength or intensity of the sense of agency one experiences when engaged in joint action. However, a second issue also arises: what form does the sense of agency take and why? That is, to what extent is agency experienced as self-agency or as joint agency, and can the sense of joint agency itself take different forms? In what follows, I consider both issues in turn. Finally, I consider emotional and motivational factors that may further modulate both the strength of the sense of agency and the form it takes.

Strength of the Sense of Agency for Joint Actions

In individual actions, the strength of the sense of agency one has for an action depends on how good the matches are between the predictions we make about outcomes at the three levels of the intentional hierarchy and actual outcomes. The same principle of congruence presumably applies for joint actions. However, as we saw in the previous two sections, in joint actions, prediction becomes a much more complex task. Agents must not just predict the consequences of their own actions at all three levels of the intention hierarchy (*self-predictions*), they must also do the same for the actions of their coagents (*other-predictions*), and finally integrate both self- and other-predictions to build predictions about the joint consequences of their combined actions (*joint predictions*). The strength of the sense of agency for the joint action (and not just one's part in it) will depend on how accurately one is able to make joint predictions, which in turn depends on the extent and accuracy of self- and other-predictions and on the manner of their integration.

One's success at making joint predictions depends on a range of cognitive abilities I described in earlier sections but also on the accessibility of relevant information. This accessibility in turn depends on the nature of the joint action. Factors such as the structure of the joint action, its scale, the degree of specialization of roles, and the longevity or transience of the collective all affect the availability of relevant information. Let us now examine these different factors.

The structure of joint actions can range from the strictly egalitarian, where all participating agents have the same degree of influence on the joint action and are equally responsible for planning it and controlling its successful execution, to the highly hierarchical, where planning, monitoring, and control are responsibilities assigned to agents high in the hierarchy. In egalitarian joint actions, the choice of the joint goal and the planning and meshing of subplans are all negotiated among the coagents, thus ensuring that they all have a relatively good knowledge of what the subplans and tasks of others are and of how they jointly contribute to the total outcome. This shared knowledge makes them well prepared to make reasonably accurate other- and joint predictions at least at the distal level. In contrast, in hierarchical joint actions, the choice of joint goals and the planning of the joint action are the concern of agents high in the hierarchy. Agents down the hierarchy typically lack detailed knowledge of the overall plans of the tasks of their co-workers. As a result of this knowledge asymmetry, agents at the top of the hierarchy, but not agents lower down, will be in a good position to make accurate other- and joint predictions.

A second important factor is scale. In small-scale joint actions, typically taking place in a shared physical environment, agents are in a position to monitor what all or most of their coagents are doing or about to do and what the consequences of their actions are, and they thus have access to the information needed to make accurate proximal other- and joint predictions. In large-scale actions, in contrast, there are too many participants for such a comprehensive monitoring to be feasible. Coagents have only (very) partial access to what others are doing and to what the joint outcomes of their actions are. To take an extreme example, think of the Allied landing in Normandy in June 1944. The individual soldier crawling on Omaha Beach in the midst of gunfire probably had very little inkling of what was going on at a broader scale and wasn't in a position to assess whether the landing as a whole was progressing satisfactorily.

A third factor to consider is the distribution of roles. In joint actions where participants have near-identical or interchangeable roles, they may have the knowledge and motor repertoire needed to precisely represent the goals and actions of their coagents and thus be in a position to make accurate proximal and motor other- and joint predictions. In joint actions, where roles are specialized and highly differentiated, this knowledge may be missing. Finally, a fourth factor that may mitigate the effects of highly differentiated roles is the stability of the association among coagents. Agents forming a long-term collective and used to acting together will typically be better able to predict the actions of their coagents and their consequences, even when roles are highly differentiated, than members of a newly formed collective.

In a nutshell, then, if the strength of the sense of agency for a joint action depends on not just self-prediction but also on other-predictions and on the joint predictions resulting from the integration of both self- and other-predictions, then participation in small-scale, egalitarian actions, with little specialization of roles and a stable group of coagents, is likely to yield a stronger sense of agency than first-time participation in a large-scale, hierarchical joint action with highly differentiated roles. Furthermore, for joint actions of the latter kind, the strength of the sense of agency experienced will depends on the position one occupies in the hierarchy. The higher up one stands in the hierarchy, the stronger the sense of agency one is likely to experience.

One should note, though, that whereas in individual actions prediction and control tend to go hand in hand, in joint actions their relationship is much less linear. In individual action, agents' predictions concern the consequences of their actions and are used to select actions, control their course, and make adjustments to them if needed. The fit between prediction and control is not perfect, and experiences of illusionary control can still arise, as shown by Wegner and colleagues (Aarts, Custers, & Wegner, 2005; Wegner, 2002, 2005; Wegner, Sparrow, & Winerman, 2004; Wegner & Wheatley, 1999), but on the whole accurate predictions tend to be reliable indicators that the agent controls the action. Thus, the more accurate they are, the stronger the sense of self-agency should be. In joint actions, however, the predictions agents need to make pertain not just to the consequences of their own actions but also to the consequences of others' actions and to their combined effects. The extent to which one might be able to predict the consequences of others' actions need not always parallel the extent to which one might be able to control their actions. As a result of this loosening of the link between prediction and control, joint actions leave much more room for spurious experiences of control.

Forms of the Sense of Agency: Self-Agency and Joint Agency

In the passage of McNeill's book that I quoted in the introduction, he describes his feelings marching and drilling on the Texas plain as involving both a sense of personal enlargement and a blurring of self-awareness and heightening of fellow feeling. His experience seems to have been simultaneously one of self-enlargement and of dissolution of the self into the collective. While there is no doubt that participation in a joint action can indeed yield this dual experience, one shouldn't hasten to conclude that self-enlargement and dissolution of the self in the collective are but the two sides of the same coin or, indeed, that one's experience when engaged in a joint action always takes the form described by McNeill. As I'll try and show later in this section, the fact McNeill had the kind of dual experience he describes may be accounted for by certain specific properties of the joint action he was participating in, properties that are not shared by all joint actions. But first, let me explain why self-enlargement shouldn't be taken as synonymous with self-dissolution.

The sense of self-agency refers to the sense one has that one is the author of an action and has control over its production and outcome. Some agents can perform actions and bring about effects that others can't. For instance, some people can wiggle their ears or play the piano while others can't. Agents may be able to do things at some stages of their life, like running or jumping, that they cannot do at earlier or later stages. Agents may be able to do things with the help of instruments that they couldn't do without. Agents may also be able to do things when empowered by relevant institutions that they couldn't do if not so empowered, like marrying couples or hiring new employees. What actions an agent can perform and what effects he or she can voluntarily bring about define what we may call the scope of the individual's self-agency, where this scope can vary from agent to agent or vary within the same agent according to age, acquired skills, available instruments, and institutional empowerments. Self-enlargement, understood as a widening of the scope of one's agency, certainly need not result in boundary loss. Rather, it appears to involve boundary expansion and, indeed, in some instances quite literally so. For example, many human and monkey studies have shown that brain representations of peripersonal space, that is, the surrounding space encompassing objects within reach, is quite plastic and that the use of tools allowing one to reach further in space results in a recoding of far space as near (Iriki et al., 1996; Farné & Làdavas, 2000).

The scope of the sense of agency should not be confused with its strength. The scope of the sense of agency refers to the range of one's action repertoire and, thus, the range of outcomes one can bring about. Its strength is linked to how accurately one is able to predict an action's outcome. In individual actions, there is typically a strong correlation between the accuracy of one's predictions and how well one controls the action and its outcome. As we saw, things are more complicated in the case of joint actions. The important point here, however, is that scope and strength are orthogonal dimensions of the sense of agency. An agent could in principle be quite limited in the range of actions he or she controls and yet predict their consequences with great accuracy and control them well, and the converse could also hold.

In many cases, joint actions allow us to bring about outcomes that a single agent could not—or could not easily—bring about on his or her own. Rousseau's stag hunt story, briefly told in *A Discourse on Inequality* (1754), illustrates the benefits of joint action. Two hunters acting together can capture a stag, whereas each hunting individually can take only a hare apiece, and a stag provides more food than two hares. Acting jointly is thus one way of widening the scope of agency. But is agency then experienced as self-agency or joint agency?

What the sense of joint agency encompasses is not easy to capture. Here's a rough attempt. In joint action, agents make their own contribution to the joint goal but must also coordinate with others (dyadic adjustments) and coordinate with others with respect to the joint goal (triadic adjustments). Contributions to the joint outcome may be important or marginal, and coordination relations can be symmetrical or asymmetrical. Roughly, then, the sense of joint agency is the sense that one's contribution to the joint outcome is commensurate to the contributions of one's coagents and that one's coordination relations with coagents are relatively symmetrical. Thus, a sense of joint agency will be fostered in situations where individual contributions are (perceived as being) of comparable importance and where coordination relations are (perceived as) symmetrical.

Other things being equal, participation in egalitarian joint actions is more likely to give rise to a sense of joint agency than participation in a hierarchically structured action. In hierarchical actions, agents high in the hierarchy can have more influence on the joint outcome than agents lower down the hierarchy, and coordination relations are highly asymmetrical with agents at the top of the hierarchy coordinating while agents down the hierarchy are being coordinated. High-ranking agents, on the one hand, are likely to experience a sense of personal enlargement, understood as an enhanced sense of self-agency rather than a sense of joint agency (conductors are famous for their inflated egos!). Low-ranking agents, on the other hand, may well experience a shrinking sense of self-agency without the compensation of a robust sense of joint agency.

In addition, very small perturbations in the relative salience of coagents can influence our perception of the importance of their contribution to the joint action. Wegner and Sparrow (2007) discuss results from social psychology experiments showing such effects. Thus, a person wearing a brightly colored shirt is more likely to be held responsible for the direction of a group discussion than someone dressed so as to blend in, even if these individuals' contributions are the same (McArthur & Post, 1977). Similarly, the physical perspective from which coactors are seen influences the perceived importance of their contribution. Looking at someone face-on rather than looking over the person's shoulder will incline us to attribute to that person a greater responsibility for the action (Taylor & Fiske, 1978). Wegner and Sparrow (2007) also report findings from their own experiments showing that small variations in the timing of action and gaze appear to influence judgments of authorship for the joint action. Thus, when two people are acting together, the person who moves first, be it by a split second, will tend to be seen as the leader of this segment of their action and will experience greater authorship of it.

We can now return to the issue left pending earlier. The experience of joint agency described by McNeill was accompanied by a blurring and dissolution of self-awareness. An experience of joint agency may also, however, leave intact one's sense of self-agency. Let us call experiences of joint agency that take the first form experiences of *we-agency* and those that take the second experiences of *shared agency*. When should we expect the experience of joint agency to take the form of we-agency rather than shared agency? To successfully coordinate their actions, coagents have to make both self-predictions and other-predictions. The more similar the actions coagents perform, the more similar their effects and the more synchronous their timing, the greater the similarity of self- and other-predictions will be and thus the harder the differentiation of self- and other-agency and the preservation of self-boundaries. The situation in which McNeil's experience of joint agency took the form of we-agency rather than shared agency presented all these features and more. The point of drilling is to get the soldiers to perform the very same actions at exactly the same time. To make self-differentiation even more difficult, the military also imposes uniform dress and hair grooming standards on their soldiers. In many joint actions, however, achieving the joint outcome requires coagents to perform coordinated yet different and complementary actions. Thus, in situations where the conditions of commensurate contributions and symmetrical coordination relations obtain and where, at the same time, coagents have differentiated roles, coagents should experience a sense of joint agency while preserving a sense of self-agency. In other words, they should enjoy a sense of shared agency.

Motivational and Emotional Factors

My focus in this chapter has been on cognitive mechanisms enabling joint action and providing cues for joint agency. Beyond the factors we have considered so far, another set of factors, motivational and socioemotional factors, may also modulate the strength of one's sense of agency and the mode in which joint agency is experienced.

There is evidence that, all else being equal, agents experience a stronger sense of agency for success than for failure and for positively valued than for negatively valued outcomes. For instance, several studies have shown that people tend to exhibit a self-serving bias in action attribution, taking credit for success but denying responsibility for failure (Miller & Ross, 1975; Whitley & Frieze, 1985). Consistent with these findings on self-serving biases in action attribution, other studies have also shown that subliminal priming of success enhances feelings of control in situations where control over the outcomes of one's actions is unclear or authorship is ambiguous (Aarts, 2007). It has also been shown that priming outcome information relatively far in advance increases experienced agency only when the outcome is linked to positive affect signals (Aarts et al., 2009). There is no reason to assume that these biases apply only to individual actions. We should expect people to experience increased agency or to be more prone to exaggerate their contributions to successful joint outcomes or joint outcomes they more positively value and to distance themselves from negative outcomes. Thus, when the French national team won the Soccer World Cup in 1998, more than a million supporters poured onto the Champs-Elysées, chanting "We won! We won!"; yet, when eight years later the French team lost to Italy in the final, the same supporters stayed at home, simply commenting "They lost!"

Beyond motivational and emotional factors common to individual and joint actions, there also appear to be factors that are specific to joint action. First, experienced agency in joint actions may also be affected by a group-serving bias, where a group-serving bias is essentially identical to a self-serving bias except that it takes place between groups rather than individuals (Taylor & Doria, 1981). Thus, the sense of joint agency would be enhanced when the joint action is successful but diminished in cases of failure. This group-serving bias may, in turn, be modulated by the strength of one's sense of affiliation with the group. Strong affiliation to members of the group could yield an increased sense of joint agency.

Second, participation in joint action may be intrinsically motivating. It has been argued that the difference between human and nonhuman joint activities lies not in the ability to read attention and intentions, an ability we share with our nearest primate relatives, but in a unique motivation to share psychological states with others, including goals, intentions, attention, and emotional states (Tomasello et al., 2005; Call, 2009).

This motivation to share intentionality appears to be manifested very early on in development. By twelve to eighteen months, infants are beginning to participate in a variety of joint actions. They are capable of helping others and are also apparently very motivated to do so, and they show some evidence of coordination of actions (Liszkowski et al., 2006; Warneken et al., 2006; Warneken & Tomasello, 2006, 2007; Carpenter, 2009). Moreover, as noted by Warneken et al. (2006) and Carpenter (2009), their behavior and emotional reactions strongly suggest that doing things together is what motivates these children and that collaborative activity is thus for them an end in itself rather than a means to achieve some individual goal. If human beings are intrinsically motivated to share intentions and to enter into collaborative activities, participation in a joint action should, ceteris paribus, be positively valued in and of itself and the experienced agency correspondingly boosted. There is also evidence (Marsh et al., 2009; Richardson et al., 2007b) that synchrony and motor entrainment lead to greater feelings of connection among coagents, fostering a sense of teamness or, as McNeill (1995) calls it, a heightening of fellow feeling. Thus, the preexisting motivation to share intentionality and engage in collaborative activities and the affiliative consequences of motor coordination may together conspire to bolster a sense of joint agency.

Motivational and emotional factors are important modulators of the sense of agency for joint actions. If the cognitive cues I discussed in the subsections "Strength of the Sense of Agency for Joint Actions" and "Forms of the Sense of Agency: Self-Agency and Joint Agency" were the only determinants of the experience of agency in joint action, we should expect a second violin in an orchestra to experience as little agency for the joint performance of the Eroica as the factory worker on the assembly line for the manufacturing of a dishwasher, as both are limited in their capacity to predict and control the total outcome of the joint activity in which they take part. The musician, however, is probably more likely to attribute high positive value to the performance of a great musical work than the factory worker to the manufacturing of a household appliance. Moreover, given the central role of synchrony in musical performances, the musician is also more likely to experience feelings of connectedness with fellow musicians than the factory worker with his or her co-workers. Similarly, if only cognitive cues mattered, the soldier subjected to close order drill should have a reduced sense of agency and not find the experience exhilarating since, apart from low-level motor control, he or she has relinquished control over his or her actions to the drill instructor. Yet, for better or worse, participation in these highly synchronous joint activities seems to foster a strong sense of socioemotional connectedness and we-agency.

Concluding Remarks

While in recent years there has been an explosion of interest among both philosophers and cognitive scientists in the phenomenology of individual actions, the phenomenology of joint actions remains to this day a largely underexplored topic. Yet, progress in our understanding of the cognitive and neural mechanisms underlying the sense of agency for individual actions as well as progress in our understanding of the cognitive processes and neurocognitive mechanisms underpinning joint actions open the road for an investigation of this topic. This chapter proposed a foray into this new territory.

Research into the sense of agency for individual actions suggests that it relies on a variety of cognitive, perceptual, and sensorimotor cues related to different levels of action specification and control and that it is governed by a principle of congruence between predicted and actual outcomes. This exploration of the phenomenology of joint action was guided by the assumption that this principle of congruence is also at work in generating the sense of agency for joint actions. However, the mechanisms of action specification and control involved in joint action are typically more complex than those present in individual actions. Thus, to understand what cues are exploited in generating a sense of agency for joint action and how the phenomenology of joint action might differ from the phenomenology of individual actions, the specific requirements that bear on joint actions and the constraints these requirements impose on action specification and control processes had to be taken into account.

In the third and fourth sections, I discussed these requirements, in particular the requirements for dyadic and triadic adjustments of intentions and actions among agents and the need for other-predictions and joint predictions they create, as well as a range of cognitive tools we use to try and meet them. I also pointed out important differences between small-scale, egalitarian actions where joint-action monitoring and control tasks are distributed among coagents and larger-scale, hierarchical joint actions where these tasks are centralized, leading to asymmetrical dependence relations among agents' intentions and actions.

In the fifth section, I explored the implications that the requirements bearing on joint actions and the various cognitive means we use to meet them have for the sense of agency we experience for joint action. I argued that insofar as the principle of congruence also applies to joint actions, the strength of the sense of agency one experiences for a joint action (rather than just one's part in it) depends on the accuracy not just of self-prediction but also of other-predictions and joint predictions. I further argued that the extent and accuracy of these predictions should be higher for agents involved in small-scale, egalitarian actions with little specialization of roles than for agents involved in hierarchical joint action with highly differentiated roles, and that for joint actions of the latter kind, it should be a function of the position the agent occupies in the hierarchy. I also argued that a sense of joint agency should be fostered in situations where individual contributions are of comparable importance and where coordination relations are symmetrical. I distinguished two modes of the sense of joint agency: we-agency, where self-boundaries and sense of self-agency appear to dissolve into the collective, and shared agency, where self-agency and joint agency are articulated rather than fused. I argued that what mode the experience of joint agency takes depends on the degree of differentiation of roles among coagents and thus on how distinguishable self-predictions are from other-predictions.

Finally, I pointed out that motivational and emotional factors may influence the sense of agency for joint actions and, indeed, that their influence may be greater on joint actions than it is on individual action. One reason for this is that beyond emotional and motivational factors common to both individual and joint actions, there appear to factors that are specific to joint action, including prior group affiliation, intrinsic motivation for shared intentionality, and the socioemotional consequences of participation in highly synchronous joint activities. These factors certainly deserve more extensive discussion than I was able to offer here. Joint action and its phenomenology constitute a domain where cognitive and motivational factors interact in highly complex ways that we need to better understand.

Acknowledgments

This work was supported by the German Ministry for Education and Science (BMBF), research cluster "Interdisciplinary Anthropology" (01GWS057). I started work on this chapter while a visiting researcher at the Department of Developmental and Comparative Psychology at the Max Planck Institute of Evolutionary Anthropology in Leipzig in 2008. I learned much from discussions with members and students of the department. Special thanks to Hannes Rakoczy, Michael Tomasello, and Felix

Warneken for many helpful comments and conversations. I am also grateful for their remarks and suggestions to the participants in the Summer Institute on Social Cognition organized by the Cognitive Sciences Institute of UQAM in Montreal in 2008 and to the participants at the Conference on Joint Attention organized by Axel Seemann at Bentley University in 2009, where I presented earlier versions of this chapter.

References

Aarts, H. (2007). Unconscious authorship ascription: The effects of success and effect-specific information priming on experienced authorship. *Journal of Experimental Social Psychology*, *43*, 119–126.

Aarts, H., Custers, R., & Marien, H. (2009). Priming and authorship ascription: When nonconscious goals turn into conscious experiences of self-agency. *Journal of Personality and Social Psychology*, *96*, 967–979.

Aarts, H., Custers, R., & Wegner, D. (2005). On the inference of personal authorship: Enhancing experienced agency by priming effect information. *Consciousness and Cognition*, *14*, 439–458.

Baker, C. L., Tenenbaum, J. B., & Saxe, R. R. (2006). Bayesian models of human action understanding. In Y. Weiss, B. Scholkopf, & J. Platt (Eds.), *Advances in neural information processing systems* (Vol. 18, pp. 99–106). Cambridge, MA: MIT Press.

Bayne, T., & Pacherie, E. (2007). Narrators and comparators: The architecture of agentive self-awareness. *Synthese*, *159*, 475–491.

Bekkering, H., de Bruijn, E. R. A., Cuijpers, R. H., Newman-Norlund, R., van Schie, H. T., & Meulenbroek, R. (2009). Joint action: Neurocognitive mechanisms supporting human interaction. *Topics in Cognitive Science*, *1*, 340–352.

Blakemore, S., & Frith, C. (2003). Self-awareness and action. *Current Opinion in Neurobiology*, *13*, 219–224.

Blakemore, S., Wolpert, D., & Frith, C. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, *1*, 635–640.

Blakemore, S., Wolpert, D., & Frith, C. (2000).Why can't you tickle yourself? *NeuroReport*, *11*, 11–16.

Brass, M., Bekkering, H., Wohlschlager, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain and Cognition*, *44*, 124–143.

Brass, M., Derrfuss, J., & von Cramon, D. Y. (2005). The inhibition of imitative and overlearned responses: A functional double dissociation. *Neuropsychologia*, *43*, 89–98.

Brass, M., Zysset, S., & von Cramon, D. Y. (2001). The inhibition of imitative response tendencies. *NeuroImage*, *14*, 1416–1423.

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

Bratman, M. E. (1993). Shared intention. *Ethics*, *104*, 97–113.

Bratman, M. E. (2009a). Shared agency. In C. Mantzavinos (Ed.), *Philosophy of the social sciences: Philosophical theory and scientific practice* (pp. 41–59). Cambridge: Cambridge University Press.

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

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience*, *13*, 400–404.

Butterfill, S. (2010). *Joint action and development*. Unpublished manuscript.

Call, J. (2009). Contrasting the social cognition of humans and apes: The shared intentionality hypothesis. *Topics in Cognitive Science*, *1*, 368–379.

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

Carr, L., Iacoboni, M., Dubeau, M.-C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 5497–5502.

Carruthers, P. (2007). The illusion of conscious will. *Synthese*, *159*, 197–213.

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

Csibra, G., & Gergely, G. (1998). The teleological origins of mentalistic action explanations: A developmental hypothesis. *Developmental Science*, *1*, 255–259.

Csibra, G., & Gergely, G. (2007). "Obsessed with goals": Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica*, *124*, 60–78.

de Vignemont, F., Tsakiris, M., & Haggard, P. (2006). Body mereology. In G. Knoblich, I. Thorton, M. Grosjean, & M. Shiffrar (Eds.), *Human body perception from the inside out* (pp. 147–170). New York: Oxford University Press.

Decety, J., & Grezes, J. (1999). Neural mechanisms subserving the perception of human actions. *Trends in Cognitive Sciences*, *3*, 172–178.

Decety, J., & Grezes, J. (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research*, *1079*, 4–14.

Dennett, D. (1992). The self as a center of narrative gravity. In F. Kessel, P. Cole, & D. Johnson (Eds.), *Self and consciousness: Multiple perspectives* (pp. 103–115). Hillsdale, NJ: Erlbaum.

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, *15*, 399–402.

Farné, A., & Làdavas, E. (2000). Dynamics size-change of hand peripersonal space following tool use. *Neuroreport*, *11*, 1645–1649.

Farrer, C., & Frith, C. D. (2002). Experiencing oneself vs. another person as being the cause of an action: The neural correlates of the experience of agency. *NeuroImage*, *15*, 596–603.

Farrer, C., Franck, N., Georgieff, N., Frith, C. D., Decety, J., & Jeannerod, M. (2003). Modulating the experience of agency: A positron emission tomography study. *NeuroImage*, *18*, 324–333.

Fogassi, L., & Gallese, V. (2002). The neural correlates of action understanding in non-human primates. In M. I. Stamenov & V. Gallese (Eds.), *Mirror neurons and the evolution of brain and language* (pp. 13–35). Amsterdam: Benjamins.

Fourneret, P., & Jeannerod, M. (1998). Limited conscious monitoring of motor performance in normal subjects. *Neuropsychologia*, *36*, 1133–1140.

Fourneret, P., de Vignemont, F., Franck, N., Slachevsky, A., Dubois, B., & Jeannerod, M. (2002). Perception of self-generated action in schizophrenia. *Cognitive Neuropsychiatry*, *7*, 139–156.

Frith, C. (2005). The self in action: Lessons from delusions of control. *Consciousness and Cognition*, *14*, 752–770.

Frith, C., Blakemore, S., & Wolpert, D. (2000a). Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society B*, *355*, 1771–1788.

Frith, C., Blakemore, S., & Wolpert, D. (2000b). Explaining the symptoms of schizophrenia: Abnormalities in the awareness of action. *Brain Research. Brain Research Reviews*, *31*, 357–363.

Gallagher, S. (2007). The natural philosophy of agency. *Philosophy Compass*, *2*, 347–357.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*, 593–609.

Gazzaniga, M., & LeDoux, J. (1978). *The integrated mind*. New York: Plenum.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences*, *7*, 287–292.

Gilbert, M. (1989). *On social facts*. New York: Routledge.

Gilbert, M. (1990). Walking together: A paradigmatic social phenomenon. *Midwest Studies in Philosophy*, *15*(1), 1–14.

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

Haggard, P., & Clark, S. (2003). Intentional action: Conscious experience and neural prediction. *Consciousness and Cognition*, *12*, 695–707.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, *5*, 382–385.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, *286*, 2526–2528.

Iriki, A., Tanaka, M., & Iwamura, Y. (1996). Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport*, *7*, 2325–2330.

Isenhower, R. W., Marsh, K. L., Carello, C., Baron, R. M., & Richardson, M. J. (2005). The specificity of intrapersonal and interpersonal affordance boundaries: Intrinsic versus absolute metrics. In H. Heft & K. L. Marsh (Eds.), *Studies in perception and action VIII: Thirteenth International Conference on Perception and Action* (pp. 54–58). Hillsdale, NJ: Erlbaum.

Jeannerod, M. (1997). *The cognitive neuroscience of action*. Oxford: Blackwell.

Jeannerod, M. (2006). *Motor cognition*. Oxford: Oxford University Press.

Jeannerod, M. (2009). The sense of agency and its disturbances in schizophrenia: A reappraisal. *Experimental Brain Research*, *192*, 527–532.

Jeannerod, M., & Pacherie, E. (2004). Agency, simulation and self-identification. *Mind & Language*, *19*(2), 113–146.

Keller, P. E., Knoblich, G., & Repp, B. H. (2007). Pianists duet better when they play with themselves: On the possible role of action simulation in synchronization. *Consciousness and Cognition*, *16*, 102–111.

Kilner, J., Vargas, C., Duval, S., Blakemore, S. J., & Sirigu, A. (2004). Motor activation prior to observation of a predicted movement. *Nature Neuroscience*, *7*, 1299–1301.

Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Current Biology*, *13*, 522–525.

Knoblich, G., & Kircher, T. T. J. (2004). Deceiving oneself about being in control: Conscious detection of changes in visuomotor coupling. *Journal of Experimental Psychology. Human Perception and Performance*, *30*, 657–666.

Knoblich, G., & Repp, B. H. (2009). Inferring agency from sound. *Cognition*, *111*, 248–262.

Knoblich, G., & Sebanz, N. (2008). Evolving intentions for social interaction: From entrainment to joint action. *Philosophical Transactions of the Royal Society B*, *363*, 2021–2031.

Knoblich, G., Stottmeister, F., & Kircher, T. (2004). Self-monitoring in patients with schizophrenia. *Psychological Medicine*, *34*, 1561–1569.

Kutz, C. (2000). Acting together. *Philosophy and Phenomenological Research*, *61*, 1–31.

Leube, D., Knoblich, G., Erb, M., Grodd, W., Bartels, M., & Kircher, T. (2003). The neural correlates of perceiving one's own movements. *NeuroImage*, *20*, 2084–2090.

Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (2006). Twelve- and 18-month-olds point to provide information for others. *Journal of Cognition and Development*, *7*, 173–187.

Marsh, K. L., Richardson, M. J., Baron, R. M., & Schmidt, R. C. (2006). Contrasting approaches to perceiving and acting with others. *Ecological Psychology*, *18*, 1–38.

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

McArthur, L. Z., & Post, D. L. (1977). Figural emphasis and person perception. *Journal of Experimental Social Psychology*, *13*, 520–535.

McNeill, W. H. (1995). *Keeping together in time: Dance and drill in human history*. Cambridge, MA: Harvard University Press.

Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin*, *82*, 213–225.

Moll, A. (1889). *Hypnotism*. London: Walter Scott.

Newman-Norlund, R., Bosga, J., Meulenbroek, R. D., & Bekkering, H. (2008). Anatomical substrates of cooperative joint action in a continuous motor task: Virtual bar lifting and balancing. *NeuroImage*, *41*, 169–177.

Newman-Norlund, R. D., Noordzij, M. L., Meulenbroek, R. G. J., & Bekkering, H. (2007a). Exploring the brain basis of joint action: Co-ordination of actions, goals and intentions. *Social Neuroscience*, *2*, 48–65.

Newman-Norlund, R. D., van Schie, H. T., van Zuijlen, A. M. J., & Bekkering, H. (2007b). The mirror neuron system is more active during complementary compared with imitative action. *Nature Neuroscience*, *10*, 817–818.

Oztop, E., Kawato, M., & Arbib, M. (2006). Mirror neurons and imitation: A computationally guided review. *Neural Networks*, *19*, 254–271.

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, *107*, 179–217.

Phillips, W., Baron-Cohen, S., & Rutter, M. (1998). Understanding intention in normal development and in autism. *British Journal of Developmental Psychology*, *16*, 337–348.

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

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

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3*, 131–141.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., et al. (1996). Localization of grasp representations in humans by PET. 1. Observation versus Execution. *Experimental Brain Research*, *111*, 246–252.

Roser, M., & Gazzaniga, M. (2004). Automatic brains-interpretive minds. *Current Directions in Psychological Science*, *13*(2), 56–59.

Roser, M., & Gazzaniga, M. (2006). *The interpreter in human psychology: The evolution of primate nervous systems*. Oxford: Elsevier.

Rousseau, J.-J. (1754/1984). *A discourse on inequality*. Trans. M. Cranston. London: Penguin Books.

Ruby, P., & Decety, J. (2001). Effect of subjective perspective taking during simulation of action: A PET investigation of agency. *Nature Neuroscience*, *4*, 546–550.

Sass, L. (1992). *Madness and modernism: Insanity in the light of modern art, literature, and thought*. New York: Basic Books.

Sato, A. (2009). Both motor prediction and conceptual congruency between preview and action-effect contribute to explicit judgment of agency. *Cognition*, *110*, 74–83.

Sato, A., & Yasuda, A. (2005). Illusion of sense of self-agency: Discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership. *Cognition*, *94*, 241–255.

Searle, J. (1990). Collective intentions and actions. In P. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in communication* (pp. 401–416). Cambridge, MA: Bradford Books, MIT Press.

Searle, J. (1995). *The construction of social reality*. New York: Free Press.

Sebanz, N., & Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in Cognitive Science*, *1*, 353–367.

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

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

Sebanz, N., Knoblich, G., Prinz, W., & Wascher, E. (2006b). Twin peaks: An ERP study of action planning and control in co-acting individuals. *Journal of Cognitive Neuroscience*, *18*, 859–870.

Sebanz, N., Rebbechi, D., Knoblich, G., Prinz, W., & Frith, C. (2007). Is it really my turn? An event-related fMRI study of task sharing. *Social Neuroscience*, *2*, 81–95.

Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2005). Listening to talking faces: Motor cortical activation during speech perception. *NeuroImage*, *25*, 76–89.

Slachevsky, A., Pillon, B., Fourneret, P., Renie, L., Levy, R., Jeannerod, M., et al. (2003). The prefrontal cortex and conscious monitoring of action—An experimental study. *Neuropsychologia*, *41*, 655–665.

Stephens, G., & Graham, G. (2000). *When self-consciousness breaks: Alien voices and inserted thoughts*. Cambridge, MA: MIT Press.

Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition*, *17*, 219–239.

Taylor, D. M., & Doria, J. R. (1981). Self-serving bias and group-serving bias in attribution. *Journal of Social Psychology*, *113*, 201–211.

Taylor, S. E., & Fiske, S. T. (1978). Salience, attention and attribution: Top of the head phenomena. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 11, pp. 249–268). New York: Academic Press.

Tollefsen, D. (2005). Let's pretend! Joint action and young children. *Philosophy of the Social Sciences*, *35*(1), 75–97.

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Hillsdale, NJ: Erlbaum.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Tomasello, M., & Carpenter, M. (2007). Shared intentionality. *Developmental Science*, *10*, 121–125.

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.

Tsakiris, M., & Haggard, P. (2005). The rubber hand illusion revisited: Visuotactile integration and self-attribution. *Journal of Experimental Psychology. Human Perception and Performance*, *31*, 80–91.

Tsakiris, M., Haggard, P., Franck, N., Mainy, N., & Sirigu, A. (2005). A specific role for efferent information in self-recognition. *Cognition*, *96*, 215–231.

Tsakiris, M., Prabhu, G., & Haggard, P. (2006). Having a body versus moving your body: How agency structures body-ownership. *Consciousness and Cognition*, *15*, 423–432.

Tuomela, R. (2005). We-intentions revisited. *Philosophical Studies*, *125*, 327–369.

Tuomela, R., & Miller, K. (1988). We-intentions. *Philosophical Studies*, *53*, 367–389.

van den Bos, E., & Jeannerod, M. (2002). Sense of body and sense of action both contribute to self-recognition. *Cognition*, *85*(2), 177–187.

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

Velleman, J. (1997). How to share an intention? *Philosophy and Phenomenological Research*, *62*, 29–50.

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *3*, 1301–1303.

Warneken, F., & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, *11*, 271–294.

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

Wegner, D. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.

Wegner, D. (2005). Who is the controller of controlled processes? In R. R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The new unconscious* (pp. 19–36). Oxford: Oxford University Press.

Wegner, D. M., & Sparrow, B. (2007). The puzzle of coaction. In D. Ross, D. Spurrett, H. Kincaid, & G. L. Stephens (Eds.), *Distributed cognition and the will: Individual volition in social context* (pp. 17–37). Cambridge, MA: MIT Press.

Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation—Sources of the experience of will. *American Psychologist*, *54*, 480–492.

Wegner, D. M., Sparrow, B., & Winerman, L. (2004). Vicarious agency: Experiencing control over the movements of others. *Journal of Personality and Social Psychology*, *86*, 838–848.

Weiller, C., Juptner, M., Fellows, S., Rijntjes, M., Leonhardt, G., Kiebel, S., et al. (1996). Brain representation of active and passive movements. *NeuroImage*, *4*, 105–110.

Whitley, B. E., Jr., & Frieze, I. H. (1985). Children's causal attributions for success and failure in achievement settings: A meta-analysis. *Journal of Educational Psychology*, *77*, 608–616.

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

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, *131*, 460–473.