Adaptive Behavior

http://adb.sagepub.com

On the Role of Social Interaction in Individual Agency

Hanne De Jaegher and Tom Froese

Adaptive Behavior 2009; 17; 444
DOI: 10.1177/1059712309343822

The online version of this article can be found at:
http://adb.sagepub.com/cgi/content/abstract/17/5/444

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
ISAB
International Society of Adaptive Behavior

Adaptive **Behavior**

On the Role of Social Interaction in Individual Agency

Hanne De Jaegher[1,2], Tom Froese[2]
[1] *University of Heidelberg, Department of General Psychiatry*
[2] *University of Sussex, Centre for Computational Neuroscience and Robotics*

Is an individual agent *constitutive of* or *constituted by* its social interactions? This question is typically not asked in the cognitive sciences, so strong is the consensus that only individual agents have constitutive efficacy. In this article we challenge this methodological solipsism and argue that interindividual relations and social context do not simply arise from the behavior of individual agents, but themselves enable and shape the individual agents on which they depend. For this, we define the notion of autonomy as both a characteristic of individual agents and of social interaction processes. We then propose a number of ways in which *interactional autonomy* can influence individuals. Then we discuss recent work in modeling on the one hand and psychological investigations on the other that support and illustrate this claim. Finally, we discuss some implications for research on social and individual agency.

Keywords agency · autonomy · cognitive gap · coordination · modeling · participatory sense-making · perceptual crossing · social interaction · social cognition

1 Introduction

Is an individual agent *constitutive of* or *constituted by* its social interactions?

Our scientific understanding of individual agency has come a long way since the inception of the cognitive sciences in the early 1970s. What started out with a focus on computational architectures in the heyday of cognitivism has been supplemented with a connectionist concern for self-organization and emergence in the 1980s, and grounded in an embodied, situated, and dynamical perspective of cognition in the 1990s. This change in our understanding, which coincides with a general shift toward privileging the concrete over the abstract (Varela, 1995), has produced some significant advances. In particular, it has become more widely acknowledged that there exists some sort of continuity between the phenomena of life and mind (e.g., Di Paolo, 2003, 2009; Stewart, 1996; Thompson, 2007; Wheeler, 1997). However, this life–mind continuity thesis has not yet culminated in a fully worked out theory of individual agency and cognition. Most theorists who subscribe to some version of this thesis recognize that an important challenge remains: how does such an account cross the gap between insect-level behavior, which has received a lot of attention since the early 1990s (see Brooks, 1991), and human-level cognition, which has been the traditional focus of mainstream cognitive science? We will refer to this particular problem of the life-mind continuity thesis as the *cognitive gap*.

Why has it been so difficult to conceive of how we can apply the insights that have worked so well for

Correspondence to: Hanne De Jaegher, Projekt DISCOS, Klinik für Allgemeine Psychiatrie, Zentrum für Psychosoziale Medizin, Universitätsklinikum Heidelberg, Voß-Str. 4, D-69115 Heidelberg, Germany. *E-mail*: h.de.jaegher@gmail.com. *Tel.*: ++44 6221 5637411, *Fax*: ++44 6221 568094

(2009), Vol 17(5): 444–460.
DOI: 10.1177/1059712309343822

models of minimal cognition (Beer, 1995, 2003) in order to understand and engineer artificial forms of the human reaches of cognition? The most important point to understand in this regard is that the cognitive sciences, even though they have undergone some radical changes, are still largely dominated by methodological solipsism. That is, after the cognitive revolution made behaviorism a thing of the past, interest in the role of the environment was replaced by an almost exclusive focus on the internal capacities of an individual cognitive agent. From this perspective the cognitive gap must indeed appear to be an insurmountable problem: how can the simple mechanisms that realize insect-level behavior be essentially similar to the ones that realize the whole spectrum of possibilities that is characteristic of human agency and cognition? Is it just a matter of increasing the complexity of our models? To be sure, it has long been shown that behavior that looks intelligent to an external observer does not have to arise from internal mechanisms that are equally intelligent (e.g., Ashby, 1960). However, the difference between insect and human cognition is so vast that a simple life–mind continuity thesis appears to be implausible (Brooks, 1997; Kirsh, 1991).

Note, however, that this extreme version of the cognitive gap is based on what can be called an "internalist" approach to cognitive science, namely the idea that agency and cognition are largely constituted by what Clark (2008) has aptly termed "brainbound" operations. To put it bluntly: if we accept the idea that our cognitive processes are essentially realized solely within our brains, then it indeed becomes impossible to conceive how the explanatory framework that has worked so well for insect-level behavior could do justice to human cognition. But what if we give up the commitment to an internalist view of the mind? This opens up the possibility that external factors could also play a constitutive role for higher-level cognition, and thus pave the way for a more moderate life–mind continuity thesis.

The idea that the environment plays a constitutive role for cognition has already been developed extensively in the cognitive sciences since the early 1990s. From robots that use "the world as its own best model" (e.g., Brooks, 1991), to dynamical accounts of how cognitive behavior emerges out of the dynamics of a brain-body-world systemic whole (e.g., Beer, 1995; Kelso, 1995; Thelen & Smith, 1994), to analyses of how environmental structures can provide "scaffolding" for cognitive problem solving (e.g., Clark, 1997, p. 45; Hutchins, 1995), and the hypothesis of extended cognition and the "extended mind" (e.g., Clark, 2008; Clark & Chalmers, 1998), the role of the environment is clearly back on the agenda. The concepts of scaffolding and the extended mind are of particular interest here because they deal with external factors that are specifically constitutive of human cognition, especially technology and language. Indeed, by arguing for the claim that human cognition is essentially a distributed phenomenon, they support a life–mind continuity thesis that is not caught in the internalist trap. However, this connection between a potential reduction of the cognitive gap and the externalism promoted by these embodied–embedded theorists has not been made explicit in their writings. The focus on technology and language, as well as human culture more generally, has obscured the fact that the principles at work in these cases can already be found in much more basic forms of agency, cognition and interindividual interaction.

In this article, accordingly, we will use the enactive approach to social cognition in order to highlight ways in which individual cognition and interindividual interactions mutually enable and constrain each other. From this perspective it will appear that these two aspects of human agency are intrinsically linked and must *both* be taken into account by researchers interested in modeling our cognitive capacities. The interaction examples discussed in this article do not rely on the specifically human capacities for language and technology, but rather on lower-level phenomena such as coordination. For this reason, they lend themselves to exploration by modelers interested in this direction of research.

We begin, in Section 2, with a brief discussion of some central concepts and their definitions: autonomy, adaptivity, agency, and sense-making. Our purpose is not to justify these definitions in their full detail—this has been done in the work we refer to. Nor do we want to say that roboticists and modelers should adhere strictly to them. One thing that will become immediately clear is that these definitions are very demanding: So demanding, in fact, that they are not applicable in the current research context. Today's robots are not autonomous in the sense of the word as we use it here. We do think, however, that practical implications can be drawn from them. These definitions can help guide research in certain directions, even if work done on their basis would seem to overshoot the current mark.

So while we think the concepts of autonomy, adaptivity, social interaction, and sense-making as defined below are not directly applicable to existing instances or models in robotics, we hope they can provide both some clarity as to the target that this research could aim at, and a set of practical sub-questions that could be addressed by the techniques of adaptive behavior research. It is the aim of this article to work out some of these implications and guidelines in the area of social interaction. Therefore, in Section 3 we review recent arguments for a natural extension of the enactive framework into the social domain in terms of participatory sense-making. In Section 4, we discuss how this opens up the notion of individual sense-making to interindividual influences. In Section 5 we discuss examples of modeling and experimental work that back up and illustrate this theoretical backdrop, which allows us to finally conclude that social agents, as well as being constitutive of the social interactions they engage in, are also constituted by them.

2 Agency and Sense-Making

Since the enactive approach to individual agency is the focus of discussion for several other contributions of this special issue (e.g., see Barandiaran, Di Paolo, & Rohde, 2009), here we will only examine definitions of the most important concepts. Later on, we use these to draw implications about the way interactional and individual elements constrain and influence each other.

One of the key concepts that ground the enactive work on autonomous agency is that of *autopoiesis*, a systemic characterization of metabolic self-production (Varela, Maturana, & Uribe, 1974). However, since autopoiesis is usually reserved for autonomy in the chemical domain we need a different concept for occurrences of autonomy in other domains. Accordingly, Varela put forward the notion of *organizational (or operational) closure* by taking “the lessons offered by the autonomy of living systems and converting them into an operational characterization of *autonomy in general*, living or otherwise” (Varela, 1979, p. 55). He says that

> autonomous systems are organizationally closed. That is, their organization is characterized by processes such that
>
> 1. the processes are related as a network, so that they recursively depend on each other in the generation and realization of the processes themselves, and
> 2. they constitute the system as a unity recognizable in the space (domain) in which the processes exist. (Varela, 1979, p. 55)

In order to separate this definition of the term *autonomy* from other typical uses found in artificial life and the cognitive sciences, we specifically distinguish it as *constitutive autonomy* (see Froese, Virgo, & Izquierdo, 2007). This concept is fundamental to the enactive approach for several reasons. First, it enables us to talk about processes of *identity generation* (see Di Paolo, 2003), that is, by being organized so as to continually create itself, the operation of the system also defines the way in which it is organized. Second, the mutuality between constitutive factors in this identity generation makes it possible for us to attribute *intrinsic teleology* to the system itself (see Weber & Varela, 2002). This means that the relations of cause and effect are at the same time relations of means and purpose. Third, because this identity is an accomplishment of the system, rather than something pre-given or imposed from the outside, it naturally finds itself in a *precarious* situation that requires continuous overcoming (see Di Paolo, 2009). In order to effectively deal with the precarious situation of their identity generation, living systems need to be able to behave in an adaptive manner. This additional property of *adaptivity* can be defined as follows:

> A system’s capacity, in some circumstances, to regulate its states and its relation to the environment with the result that, if the states are sufficiently close to the boundary of viability,
>
> 1. tendencies are distinguished and acted upon depending on whether the states will approach or recede from the boundary and, as a consequence,
> 2. tendencies of the first kind are moved closer to or transformed into tendencies of the second and so future states are prevented from reaching the boundary with an outward velocity. (Di Paolo, 2005, p. 438)

Thus while self-constitution establishes a purpose and a perspective, adaptivity is needed so that the system can anticipate and act upon that which it needs for its self-constitution and that which may threaten it.

Systems that are both autopoietic (or, more generally, constitutively autonomous) and adaptive are capable of *sense-making* (see Di Paolo, 2005; Weber & Varela, 2002). Sense-making is the capacity for a system to enact a world of meaning.[1] This entails that the system is able to interact with the environment in terms of the consequences that its interactions have for the conservation of its identity. In other words, such consequences have significance or value for the system, namely in relation to the processes of its identity generation. This is because self-generation is a source of norms: different events will contribute differently to its continuation, some enhancing it, others putting it in danger.

This self-given normativity applies at the metabolic level and all the way up to the highest kinds of identities that human beings can make for themselves. At the direct metabolic level this means that certain elements of the environment count as nutrients and others do not. But it can also mean that through adaptivity the system is able to make sense of more indirect forms of coupling with the world and sustain the requirements of several organic norms. Such is the case for migrating species that travel long distances to find nourishment when, for instance, breeding grounds are far away from the main source of food, enduring long periods of hunger and navigating through difficult landscapes in the mean time.

In this view, the idea of sense-making is intimately connected to the conservation of a precarious identity. Metabolism is the most basic example, but in a given organism it is possible for several such processes to overlap, bringing forth the possibility of non-metabolic norms such as habits of behavior or sustained forms of relations within a group. In the human ranges, norms of interaction with the world can go so far as to put severe constraints on metabolic identity, for instance in extreme dieting. A social identity also needs to be preserved. When a person risks getting fired from her job, her identity as an employee is in danger of breaking down, but this does not necessarily have an immediate effect on her metabolism.

An important aspect of sense-making is the central role of *movement* in it. As Sheets-Johnstone (1999) says, agents make sense of the world in movement. Even though Sheets-Johnstone's approach is primarily phenomenological and in the first place applicable to humans, this principle is well-known in adaptive behavior research. For instance, Beer (2003) describes an evolved agent capable of discriminating objects in terms of their shape through active scanning, as opposed to some kind of internal reconstruction of the sensory pattern followed by appropriate calculations to generate the categorization. Thus, the process of artificial evolution found a solution to the task that uses a pattern of moving exploration. While this agent is not autonomous, its behavior illustrates the role of movement in aspects of sense-making (for the role of moving exploration in cognition, see also O'Regan & Noë, 2001).

In short, autonomous adaptive systems enact a world of meaning and value through their movement in it. Sense-making, then, is the capacity of a system to enact a world and imbue it with significance from its own point of view. An autonomous adaptive system does this in relation to the processes of identity generation that realize it as the particular organism-environment systemic whole that it is. We can now define an *agent* as any autonomous system capable of sense-making in its interactive domain (see Barandiaran et al., 2009).

This definition of agency does not tally with the way the notion is used in most of robotics research today (see Froese & Ziemke, 2009). According to the picture of autonomy, adaptivity and sense-making sketched here, no robot with these "qualities" exists as yet. Robot behavior can conform to an externally imposed norm, but robots still do not produce norms themselves (Di Paolo, 2003) and this is currently the main obstacle for calling these robots agents in the sense of the word espoused above. Nevertheless, robotic modeling can contribute to our understanding of these concepts by studying specific sub-aspects of these phenomena, without necessarily having to replicate them in their entirety. For instance, it is possible to model phenomena such as social interaction even if agency in the sense above cannot yet be modeled, as we will see in Section 5.[2] But before showing practical examples, we discuss a theoretical extension of the enactive ideas into the social domain in order to get a first grasp of how agency and sense-making might be affected by interindividual interactions.

3 Social Interaction and Participatory Sense-Making

The notion of sense-making has recently been extended into the domain of intersubjectivity in the form of *participatory sense-making* (De Jaegher & Di Paolo, 2007). The idea is that participants to a social encounter can participate in each other's sense-making. The proposal emphasizes the interplay between two levels of autonomy: that of the interaction process as such—a non-intuitive idea that we will examine more below—and that of the individuals engaged in interactions. While social interactions are to an extent grounded in individual agency and sense-making, at the same time they also constitute novel domains of possibilities for sense-making. Interactions in this view are irreducible to the individual level. Moreover, interactions *as such* can influence the individuals (over and above the influences that interactors can exert on each other—a more generally accepted idea). According to the participatory sense-making proposal, understanding social cognition must start from the *process of interacting* that participants in a social encounter routinely engage in. The notions of coordination, the autonomy of the interaction process and sense-making help to operationalize this idea.

Coordination is defined by De Jaegher and Di Paolo (2007, p. 490) as "the non-accidental correlation between the behaviors of two or more systems that are in sustained coupling, or have been coupled in the past, or have been coupled to another, common, system." Coordination often happens in physical and biological systems and has been heavily studied in these two areas (Buck & Buck, 1976; Kelso, 1995; Winfree, 2001). The fact that it is ubiquitous in these domains suggests that it is a process for which high-level cognitive mechanisms are not necessary. It has also been studied in neuro-behavioral phenomena (de Rugy, Salesse, Oullier, & Temprado, 2006; Fink, Kelso, Jirsa, & de Guzman, 2000; Jirsa, Fuchs, & Kelso, 1998), and here too it seems that no high-level cognition is necessarily involved. Coordination can be achieved by biomechanical means. It has also been found in interpersonal settings such as conversation; an entire field of study in the social sciences is devoted to the study of interpersonal synchrony (e.g., see Goffman, 1983; Goodwin, 1981; Grammer, Kruck, & Magnusson, 1998; Sacks, 1992). A study by Schmidt and O'Brien (1997) found that coordination can even happen between people when they are explicitly instructed *not* to coordinate.

Coordination is not an on-off process, it is possible to move "into" and "out of" coordination. To understand this, the notions of relative and absolute coordination, originally coined by von Holst, are helpful (Kelso, 1995, p. 98). Kelso explains how coupled systems can move into and out of zones of absolute coordination through regions of relative coordination. Coordination has many possible ways of manifesting: behaviors can coordinate in time (think of synchrony), they can be imitated, mirrored, anticipated, and so on. Coordination can break down. Instances of coordination breakdown are of significance for this approach to the study of social cognition because when coordination breaks down, possibilities for reestablishing it open up. Depending on whether reestablishment happens or not, the interaction can continue or end, respectively.

De Jaegher & Di Paolo, 2007 claim that the interaction process itself can become autonomous. This idea can be illustrated by exchanging the relevant terms in Varela's definition of autonomy already cited above with terms relevant to the social interaction. Thus, we can say that

> the interaction process is organizationally closed. That is, its organization is characterized by processes such that
>
> 1. processes of interindividual coordination are related as a network, so that they recursively depend on each other in the generation and realization of the processes themselves, and
> 2. they constitute the interaction process as a unity recognizable in the space (domain) of relational dynamics.

In order to clarify this, let us take a look at the interplay between interaction and coordination. Interindividual coordinations between movements (including utterances, see Gallagher, 2005) can lead to the emergence and then maintenance of an interaction process. A good example of this is when an interaction process emerges between people who did not intend it to happen. For instance, when two people approach each other from opposite directions in a narrow corridor, they sometimes end up in a tangle of movements instead of simply walking past each other, even if their individual goal was to do the latter. What makes this interaction process emerge and briefly subsist is the coordination between the movements of both people — it could be a coordination by mirroring as well as a temporal one: stepping to the same side at the same time. As soon as this happens, they enter into an interaction. Because of their initial coordination, they end up exactly in front of each other and thus in each other's way again. What happens next is likely to be more coordination. Maybe they will each step to the other side and repeat the mirroring moves a couple of times on each side. Interaction and coordination can thus sustain and nurture each other, resulting in a self-sustaining process of continuing interaction, thus illustrating point 1 of the definition. This interaction is a particular one and as such recognizable in the space of relation dynamics (point 2). It is an interaction that emerges there and then, different from all other such interactions (no two of them are exactly the same), while at the same time sharing many general characteristics with other interactions between people (De Jaegher & Di Paolo, 2007).

The corridor example also shows that an interaction process, in becoming autonomous, can override the intentions of the individuals involved in it. As long as the interaction process continues to unfold in this way, the intention of each individual to just walk past the other is thwarted while, at the same time, new intentions can form.

The idea of the autonomy of the social interaction process serves as a starting point for explaining how social understanding happens, namely in the "in-between" between people. The proposal aims to address the criticisms made of traditional, individualistic approaches that they wrongly assume that social cognition is a capacity confined to individual reasoning or simulation capacities, and that they do not take into account interpersonal engagement (see Gallagher, 2001; Hobson, 1991; Reddy & Morris, 2004). The argument goes like this: if there can be interpersonal coordinations of movements (including utterances), and movements play a central role in individuals' sense-making activities, then these sense-making activities themselves may get coordinated in interaction. According to the characterization of coordination given above, this does not require higher-level cognitive mechanisms. The coordination of sense-making activities is called *participatory sense-making*. There is a spectrum of interpersonal meaning-making ranging from guiding (e.g., pointing something out to one's interaction partner and thereby orienting his attention and or understanding) at one end, to truly joint sense-making (the generation and or transformation of meanings in interaction, where this formation or transformation cannot be attributed to either of the interaction partners alone) at the other end. Participatory sense-making opens up domains of sense-making that are not available to an individual alone. Social cognition then is not an application of a general cognitive problem-solving capacity to a specific domain (the social one), as traditional approaches assume it is. Rather, it is an agent's engagement in a certain kind of embodied and situated interaction with another agent. This engagement is characterized by the autonomy of the interaction process, established in the co-regulation of the process by the two agents, who each maintain their own autonomy during the process. The interaction as such thus plays a role in the interpersonal generation and transformation of meaning.

It is not our aim to exhaust the meaning aspect of participatory sense-making (but see for instance De Jaegher, 2009). Our main claim here is that the interaction process can also shape the *constitution* of the individual agents—in other words, that their constitutive autonomy can be changed by the interactions they engage in. The next section is devoted to this argument.

4 Impact of Social Interaction on Individual Agency

We now examine the interplay between the autonomy of the interaction process and the respective autonomies of the individuals engaged in it, in order to derive consequences for understanding individual agency.

What is interactional autonomy? In 2007, De Jaegher and Di Paolo said that

> [s]ocial interaction is the regulated coupling between at least two autonomous agents, where the regulation is aimed at aspects of the coupling itself so that it constitutes an emergent autonomous organization in the domain of relational dynamics, without destroying in the process the autonomy of the agents involved (though the latter's scope can be augmented or reduced) (p. 493).

This definition attests to the autonomy of the interaction as a process in the domain of relational dynamics.

We have illustrated (with the corridor example) how this happens, and below we will discuss some models and experiments that also support this idea. This view, however, also gives a starring role to the individuals, through their autonomies. What is special about a *social* interaction (as opposed to a non-social one, e.g., an interaction between agent and physical world) is the interplay between these specific autonomies.

We have suggested above that an interaction can be autonomous in the sense given by Varela's definition and, furthermore, that interaction processes can self-generate. For instance, even in situations where neither of the interactors want to interact, an interaction can come into existence and self-maintain for a while. This kind of situation is a serious "precarious condition" for the interaction's existence. What makes an interaction emerge and self-sustain? What are the material substrates for this self-production? A social interaction emerges upon the getting-together of two autonomous, social agents. Is there something about this kind of encounter that entails the possibility that an autonomous interaction process can emerge out of it? All that is required for coordination is a degree of similarity between the elements to be coordinated, and that they "behave" in a very basic way (systems that do not move or otherwise behave are unlikely to get into any kind of coordination). Both of these are true of social agents, especially if the ones under consideration share sufficient similarity, for example, humans or other animals. Coordination of social agents' movements can happen; and because animate movements play such a role in individual sense-making activities, as we have seen, these sense-makings themselves can coordinate. This makes it possible for individual social agents to participate in each other's sense-making, without a need for high-level cognitive capacities for the coordination to take place. What are the implications of this view for the constitution of individual agency?

In the following, we list a range of implications. We intend this more as an exploratory move (rather than a deep analysis of each and their interrelations), to help understand where modeling and experimental research in this area could go. We hope the descriptions that follow may serve as an inspiration for new models and experiments.

4.1 Regulation of Interactions Is Not Completely Down to Individuals

Who or what regulates the coupling in a social encounter? In the corridor situation, two agents encounter each other and end up interacting despite neither of them having the intention to do so. As soon as each steps to the same side at the same time, interaction and coordination processes unfold. As they continue to step to the other side at exactly the same time, interaction and coordination mutually reinforce each other, and the stepping to the side repeats itself a few times. What regulates the coupling here? We suggest that it is the coordination process itself, or rather, the mutual reinforcement of interaction and coordination.

As long as the interaction is unintentional, the individuals do not fully control it. But of course each agent has the possibility of starting to regulate the interaction at any point, and this, in fact, is how people often get out of these kinds of entanglements. They can, for example, break the coordination of movements and start to regulate at another level by for instance speaking ("after you"), stopping, or laughing. What happens here is an intended breakdown of the coordination, in order to interrupt the interaction pattern and for each individual to continue with their own business. A willful breaking of the coordination pattern like this allows individuals to regulate the interaction in order to end it.

Thus, social interactions can run their course without any of the individuals involved in it having strict control over the process, although the individuals may at each point (try to) regulate it. There may also be external factors that contribute to the regulation of the interaction, for instance technological mediations. The interaction process forms in the domain of relational dynamics.

4.2 Intentions Can Be Constructed in Interaction

Two individuals in an interaction such as the corridor encounter are unable to follow their own goal for a brief moment. But while they are "trapped" in the process, they may develop new intentions. In this way, the unintended interaction may, for instance, turn into a desired one. A scene in Hitchcock's *North by Northwest* beautifully illustrates this. Cary Grant's character Roger Thornhill, while on the run from the police on a train, bumps into the beautiful Eve Kendall (Eva Marie Saint). A corridor situation emerges between the two and, for a couple of seconds, they are engaged in a kind of coordination dance. In the end, Thornhill steps into the compartment that Kendall had come out of. He had wanted to run through the carriage as fast as he could, away from the police. His bumping into the lady and being held up by her was an unintended mishap. However, as the "bumping into each other" unfolds, something else becomes clear: maybe they *do* want to meet…

Precisely because not everything is down to the individuals, the course of the interaction may provoke them into realizing (explicitly or implicitly) new intentions—or old but latent ones—that may become salient because of the unfolding of the interaction. Gibbs (2001) gives some more examples of this.

4.3 Interactional Breakdowns and Recoveries, and the Emergence of Meaning

A related issue is the fact that new meanings can emerge in interaction. The interactional autonomy does not necessarily override individual intentions all of the time. The interactional autonomy, like coordination, is not a matter of strictly on or off, but rather one that fluctuates. It is in these fluctuations of the regulation of the process that meanings can arise or be transformed. After all, the fluctuations are where the individuals' roles in the process can be subject to change. Meaning, of course, can only be made by agents (in contrast to the autonomy of agents, the autonomy of the interaction does not seem to imply that it itself can have a perspective of significance), and the points at which these changes in the role of the individual happen seem the most fertile ground for meaning formations and transformations. When a pattern of interaction breaks down, any co-adaptation that leads the participants to regain a coordinated situation can be the source of a new significance about their own actions, each other's, and the world more generally. If a new pattern emerges, and sediments, it can become part of the future repertoire of this particular set of interactors.

Is the point here that the interaction autonomously generates a new significance? Of course not. It is rather that interactions influence their participants' understandings. This may seem a trivial point, but it is not, for its implications are non-traditional. No individual in a social situation is always and totally responsible for all that happens, not even in turns. Some of the "responsibility" can quite literally lie outside of all the participants to a particular interaction. However, this does not detract from the fact that they can influence interactions. They can for instance mend faulty interactions or damage good ones by repairing or breaking the process.[3]

4.4 Actions Can Be Resolved or Completed in Interaction

Certain actions may only be completed in interaction. The act of giving is a good example of this. An agent may move the hand in which he is holding an object in the direction of his interaction partner, while not yet sure whether he is intending to give it or not (he could still just show it). Sometimes the act of giving is only fully established (rounded up, so to speak) as the other reaches out and grabs the object (Fogel, 1993, describes an example like this in his book, pp. 20–21; see also De Jaegher & Di Paolo, 2008). Also, actions that are normally outside the scope of an individual agent may be possible in the context of an interaction.

4.5 New Skills Can Be Acquired in Interaction: Among Them, the Skill of Interacting Itself

One consequence of the previous three points is that engaging in interaction can provide the agent with new skills. An example of this is the transition from grasping to pointing in infant development. Vygotsky (1978) believed that pointing develops out of grasping. We can imagine how a recurring kind of interaction in the child's everyday life can make this happen. The infant reaches out, intending to grasp a toy but unable to reach it. Sometimes, this action may be completed by an attentive caregiver. If this happens repeatedly, this sort of event may ground a new significance of the gesture of reaching which may eventually become one for asking, and later referential pointing (see Fuchs & De Jaegher, in press). Thus the infant acquires, through repeated interactional engagement, a new skill: that of pointing. This particular example shows how the skills acquired in interaction can be redeployed as new forms of participation. Other social skills may develop in similar ways.

4.6 The Scope of Individual Autonomy Can Be Altered During Interaction

An individual's autonomy can be reduced or enlarged in interaction. Two people may intend to have a pleasant, friendly interaction, but the interaction pattern may prevent them from doing so. This can happen in interactions with a certain history, that is, between people who have interacted before. An interesting example of this is Granic's (2000) study of aggressive interactions between parents and teenagers. She showed that such interaction patterns can get entrenched and become very hard to break, while having a tremendous effect on the participants. Breaking such patterns does not depend much on the individuals per se, which is why they can be so hard to dissolve and may require intervention from outside the situation. In cases like this, the autonomy of each interaction partner is diminished, in the sense that it is hard for them to break the pattern and change the direction in which the interaction is going.

Many other examples of this can be found. One lies in a comparison between different styles of teaching (see e.g., Lave & Wenger, 1991). In some old-fashioned views on the teacher–pupil relation, the student's role was quite limited. He would be seen as not much more than a receptacle to be filled with the knowledge that the teacher already had. Hence, for instance, the popular practice of learning by rote. In such learning, the student is not supposed to have a very active attitude toward the material. The idea here is that of—literally—a transfer of knowledge, from one container to another. In contrast, more progressive approaches to teaching demand of the learner an active and pro-active role in assimilating and accommodating the material. Here, the focus is on participation in a process of producing knowledge. This has a big impact on the relationship between teacher and student. In the first approach, the student is given a limited role, he is often merely made to listen. This can feel, to the student, like a stringent reduction of his autonomy (sometimes leading to strong adverse reaction), and it really is that too: his influence on the course of the teacher–student interaction is limited to receiving the knowledge. In the second kind of approach, the role of both student and teacher changes. The student has more autonomy and can influence the interaction process more, while the teacher is now also a more active participant in the teaching process. She can expect more engagement from her student, which can include questions, criticisms and so on. This requires a more active engagement from her in return. It also entails that a greater responsibility for his own learning lies with the student himself.

Thus individual participation in interaction can be enhanced or limited, for a number of reasons: customs, habits, social roles, inequality between the participants (e.g., in defendant–inquisitor situations), established views on what ought and ought not to be done, and so on. And one of the driving forces for this can also be the interaction process itself.

4.7 Engaging in a Social Interaction Implies Reciprocity

One of the characteristics of agency is the fact that agents regulate their world. This implies a certain asymmetry: the physical world typically does not regulate you back (see Barandiaran et al., 2009). This is different in interactions with other agents, who do (attempt to) regulate you back. However, this does not necessarily lead to symmetry. In fact, as we have seen so far, one thing to note about inter-agent interactions is that they are subject to many different kinds of tensions. It is precisely these tensions that make the study of the social interaction so interesting in relation to understanding social understanding. In a social interaction, each agent is at once prodder and prodded, and so is his interaction partner. What does it mean for participants to engage in each other's sense-making activities? It means to influence the other's meaning-making activities *as well as* to be influenced by theirs. This happens at the level of movements (e.g., take a few tentative steps to the right while strolling through town with a friend) and at the level of intentions ("ah yes, good idea, let's go and have a look over there").

The crux is in the kind of relation, which includes the kind of influence that can be exerted. In social interactions, the possibility for influence is always reciprocal. This is a definitional aspect of social interactions. If it were absent, we would no longer be in the presence of a social interaction. In social interactions, while changing myself (taking a different perspective, making-sense in some other way), I am also influencing another. And while influencing another, I am changing myself. The way this works in each particular interaction can add a specific possibility to my autonomy, or detract from it—different interactional roles and identities become possible.

4.8 Interaction Allows for Delegation

The work involved in the fulfillment of individual goals can be (partly) transferred to the interaction/interactive action (i.e., you do not need to be able to do everything yourself). In interactions, it is possible to shift tasks, responsibility, and so on, onto the other, and even sometimes onto the process. As we will see below in the example of perceptual crossing (see the experiment described in Section 5), some tasks can only be fulfilled when done in interaction. This may lighten the load of the task for individuals. This process may also be seen in therapy, or even while sharing difficult times with a friend. There is work to be done on understanding how this happens.

4.9 Social Interaction May Lead to a New Form of Individuality

At the extreme end of social agenthood, for instance in humans, where social interactions are so prevalent and intersubjectivity can be so deeply meaningful, individuals on their own are not purely individual, but always to an extent social. Examples are: talking to oneself, incorporation of social norms to regulate one's individual activities, and even language-mediated conceptual thought (see Vygotsky, 1986). In other words, social agents end up being so social that even when alone, they are not non-social individuals—in the sense that their social embeddedness continues to influence their behavior to a large extent.

4.10 The Individual's Role in Interaction

All this attention to the autonomy of the interaction should not make us forget the important and indispensable role individuals play in interactions. It is important to recognize that, even though a great deal depends on the interaction process, the individuals are not ignored in this proposal. On the contrary, the emphasis on the interaction process gives the individual a place in intersubjectivity that he did not have before. If the interaction can become autonomous, the individual's own autonomy needs to be taken especially seriously.

Thus, the participatory sense-making proposal has potential implications for our understanding of individual agency. According to the definition of social interaction, individual autonomy should not break down under the interaction. If it did, the interaction would stop being a social one. This, of course, is not just a definitional aspect. It is clear that individuals influence the interaction. But can we expect individual agents to be influenced in their agency and autonomy by interactions? It seems clear that things are possible in an interaction that are not possible for an agent on their own. This is a straightforward effect that the interaction has on individual agency. Moreover, what an agent can do in interaction can change him in such a way as to make different things possible outside of that interaction as well. Social agents acquire new skills, intentions, and understandings in social interactions. If we want to advance on modeling aspects of human agency, it will be good to take these suggestions into account.

The reader will have noticed that common threads run through all of the implications we have listed. In a way they cannot be as neatly divided as we have just attempted. Intentions, roles, skills, goals, identities, autonomies, significances, and so on, are in fact not so easily unraveled and separated from each other. In autonomous agents, there will always be an extent to which these aspects are entangled in each other. The purpose of pulling them apart in the way we have done here is to make this proposal more amenable to modeling and experimenting. But whatever model or experimental set-up is designed, if the larger goal is to understand human agency, we must keep in mind the complexity of how it all fits together. If we lose sight of this, our models may end up on detours that lead us too far astray from the overarching objective.

5 Examples and Models

In this section, finally, we discuss two pieces of work that back up and illustrate the theory, as well as exemplify the kinds of aspects of these phenomena that can be researched using modeling and experiments. The investigations we describe here attest (a) to the enactive idea that the interaction can take on an autonomy, and (b) to the role of the interaction process in social tasks (which are typically only thought of in exclusively individual terms).

We start with a model. Simulation models that are generated by a method of evolutionary robotics (see

Beer, 2003; Harvey, Di Paolo, Wood, Quinn, & Tuci, 2005) are an integral part of the enactive cognitive scientist's toolbox. Because they incorporate only essential aspects of the behavior of interest, they are simple enough to study, yet can provide us with an improved understanding of complex phenomena that involve emergent, many-layered, causally spread, non-linear dynamics (Di Paolo, Rohde, & De Jaegher, in press). Simple forms of social coordination have been the target of several models in adaptive behavior research using simulated or real mobile agents. These models study the evolution of communication or coordinated behavior under situations that promote agents to engage in collaborative interactions or under situations of conflict (e.g., see Di Paolo, 2000; Marocco & Nolfi, 2007; Quinn & Noble, 2001). Many of these models investigate minimal forms of engagement that support some of the ideas we propose here, others are complementary.

Here we discuss the insights that Froese and Di Paolo (2008) obtained by implementing a minimalist simulation model of Murray and Trevarthen's (1985) double TV monitor experiment. In this psychological study 2-month-old infants were animated by their mothers to engage in coordination via a live double video link—mother and infant are each in a different room and interact with each other via a TV screen. However, when the live video of the mother is replaced with a video playback of her actions recorded previously, the infant becomes distressed or removed. These results, and those of a more rigorous follow-up study by Nadel, Carchon, Kervella, Marcelli, and Réserbat-Plantey (1999), indicate that 2-month-olds are sensitive to social contingency, that is, to the mutual responsiveness during an ongoing interaction, and that this sensitivity plays a fundamental role in the unfolding of coordination. Traditional explanations of this sensitivity have focused on innate abilities. For example, Gergely and Watson (1999) have postulated the presence of an inborn cognitive module which enables the detection of social contingency, and Russell (1996) hypothesizes that infants have an innate capacity to understand intentionality and to process agency. Are these postulations of innate capacities on the part of the infant necessary in order to explain the empirical results?

Froese and Di Paolo's (2008) study found that stable and robust coordination can be reliably established between simulated agents. The goal of the agents is t]TJ0actifocnso1.2a 0asturated bliar[(impleTd[(oof thnd inte)6(n55]TJ-0.00

Recently, Auvray, Lenay, and Stewart (2009) performed a psychology experiment in which they asked pairs of human individuals to indicate when they think they encounter each other in a thoroughly stripped-down virtual environment. Participants are located in separate rooms, blindfolded, given only a computer mouse with which they can move left and right on the table top, and a tactile stimulus device that taps the tip of a finger on their other hand. While moving the mouse left and right, participants can encounter objects in their virtual world. Each time they do so, they receive a tap. There are three objects that each participant can come across in their world: a fixed object and two moving ones. The taps received at the time of encounter are of the same quality (duration and strength) for each kind of object. The first task for participants is a training task: to distinguish between fixed and moving objects. This is easily done since "encounters" with each of these categories of objects generate a different pattern of finger-taps-in-relation-to-movement: the fixed object generates a tap each time the subject moves over it—in the same place; while the moving objects also generate a tap each time the subject moves over it, but—as the object moves around—the participant's mouse will be in a different location each time a new tap comes.

After this familiarization task, the participants are asked to click whenever they think they encounter the other. The subjects do not know which of the moving objects represents the other. In fact, unbeknownst to them, one of the moving objects represents the other person's avatar, while the other one represents the other person's "shadow," an object attached to the avatar at a set distance from and identical to it. Since shadow and avatar are tied to each other, they move in exactly the same way, which makes the task difficult to solve, in fact, at first sight almost impossible to solve. The experiment is dubbed an investigation of "perceptual crossing" because even though they are not aware of it, what the participants are really asked to do is to distinguish when they are in the presence of the "perceiving" part of the other (his avatar) from when they are in the presence of his "shadow" (with which the other does not perceive anything) by clicking in the first case only.

Participants did manage to solve this task. How? A count of the participants' individual responses (clicks) showed that they were equally likely to click on the avatar and the shadow. The probability of clicking on avatar versus on shadow were statistically indistinguishable. In other words, individually the participants were not able to distinguish between them. So what did the trick? The solution lies in the fact that *both* participants are trying to solve the task. This creates a dynamic whereby when one bumps into the other's shadow, the other does not perceive being perceived, and will not stay where he is for long (he is also searching). This gives the shadow less exposure, so to speak. The shadows are encountered less often than the avatar. This result was discovered by looking at the interaction as a temporally extended whole. It was the collective dynamics of the interaction that allowed the participants to solve the task, because, overall, they encountered the avatar more often. Only when both participants run into each other with their "sensory parts" does the dynamic of interaction enter into a globally stable state. The solution to this task, therefore, was not based on any individual recognition capacity, but on a global dynamics of the interaction.

This is a task and set-up for which there *is* no individual solution, only an interactional one. The experiment is interesting in that it shows that participants can indeed spontaneously interactionally (and *only* so in this case) solve a task, without being aware that this is the case. Research on physical cooperation shows a similar result, with the difference that the task in their case could have been solved individually (Reed, Peshkin, Hartmann, Colgate, & Patton 2005; Reed, Peshkin, Hartmann, Grabowecky, Patton, & Vishton, 2006). They were able to show, however, that performing the task (moving a crank in order to move a virtual object into a target position) benefited from being done in pairs. It was performed more efficiently and quickly that way, even if the subjects thought they did better individually.

The simplicity of Auvray et al.'s experimental set-up makes it very amenable to minimal cognition modeling. This task was undertaken by Di Paolo, Rohde, and Iizuka (2008), who evolved simple agents controlled to solve the same task. Their model confirms the explanation provided to account for the empirical results, but goes further by suggesting novel hypotheses (for instance, the potential role of sensorimotor delays) and unexpected correlations that were later confirmed on the original data (such as participants sometimes getting stuck on static objects). One interesting finding is how the agents resolved the task. The same mechanisms as in the human case seem to be at work for the distinction between the avatar and the shadow. Agents may scan the other's shadow, but since the other is still searching, the situation is not as stable as when both agents are scanning each other's sensors. But unlike humans who have complex proprioceptive capacities and spatial awareness, it is rather hard for the agents to distinguish between a coordinated oscillatory crossing between each other and the oscillatory scanning of the static object. Agents in the model are actually seen to scan static objects for a long time and then to move on. While they do this, their sensorimotor patterns are very similar to the situation when they are scanning each other. How do they distinguish between the two situations? They use a clever trick. The apparent size of a scanned object depends on the speed of the agent (slower scans take longer time on the object and produce longer sensor activations). Mutual crossings of two agents in antiphase effectively halves the time of sensory stimulation. This is the sensory factor used by the agents to decide to stay in contact with each other. Crucially, the perceived size of the object is in this case a "social construct": it is systematically affected by the coordinated movements to create a perceptual difference that results in a specific course of action that in turns keeps sustaining the coordinated movements. We see clearly in this concrete example what we have tried to indicate in more general terms as the autonomy of the interaction.

Other models have also contributed to an exploration of some of the theoretical issues raised in this article. In general these are models that take a minimal form. For instance, Quinn (2001) has shown the evolution of a minimal form of coordination between simulated mobile agents based on what might externally be construed as an emergent signal that leads them to allocate roles and allow them to move together (their set objective). Because models such as these tend not to assume a complex substrate for communication between agents nor a complex cognitive apparatus (in fact, Quinn's models and the models described above only provide agents with very simple distance or presence sensors and a very small recurrent neural network), they open the possibility for exploring how much can be achieved in the interaction process itself and often lead to unexpected revelations about the constitutive role of the interaction in the agents' performance and capacities. This is in contrast with other approaches to modeling the evolution of communication in artificial agents (see e.g., Cangelosi, 2007; Steels & Kaplan, 2002), where crucial cognitive capabilities (to categorize perceptions, to understand a pointing gesture, to achieve joint attention, etc.) are assumed to exist (and thus pre-programmed). Such crucial individual capacities, as we have suggested in Section 4, are in fact often to the result of interaction and might indeed be constituted and sustained interactionally. To assume them as pre-existent or to isolate them from change resulting from interactional experience risks framing the results within the individualistic mindset that we have criticized here.

6 Upshot

We set out to investigate the question whether agency is *constitutive of* social interaction or *constituted by* it. The motivation behind this question runs in parallel to the question of whether the enactive approach, which at first sight may seem to over-specialize in low-level forms of cognition, can be applicable all the way to full-blown human-level cognition. We investigated the idea that this could be done via the study of agents' social interactions. Starting from the idea that *social interactions can take on an autonomy of themselves*, we derived ways in which the constitution of individual agency can be influenced by the interactional contexts in which social agents routinely find themselves. The ways in which this can happen include: an agent's individual intentions, sense-making, and social skills can be affected by social interactions; agents can, through co-regulation, expand the domain of their actions; in some cases, being skilful at social interaction can create a new form of individuality; social encounters, unlike interactions with the physical world, require and engender reciprocity. With regard to the question we started out with, we can now say that individual agency, at least individual human agency, is to a large extent determined by social factors. If this is the case, further investigations in this direction could be one way to advance on connecting low-level cognition and higher forms of mental life.[4]

We took the idea of participatory sense-making as our theoretical starting point. This proposal emphatically roots social cognition in the dynamics of embodied coordination and interaction between autonomous agents. This makes it amenable to simulation and modeling research, as we have illustrated here. When we started from the very demanding definitions of agency, autonomy, and sense-making in Section 2, our idea was not to move the goalposts of empirical research on these issues past the current horizon. Rather, it was to provide direction and inspiration without risking losing sight of the larger objective of this kind of research. We chose certain modeling and experimental examples to illustrate that the autonomy of the interaction process and its role in social tasks and in the constitution of individual agency can be studied, especially if we focus on the right aspects of these phenomena. These aspects centre on how agents coordinate and what role the interaction as such plays in this coordination. This, in turn, allows us to come back to the theory to sharpen and enrich it, because modeling and experiments can be used as tools to investigate the causal efficacy of the interaction process in a non-mysterian manner (see Boden, 2006).

As we have said, this work is intended as an exploration. We have hoped to do three things: (a) sketch how the participatory sense-making proposal relates to and perhaps can inspire further robotic modeling and psychological experiments, and (b), in doing this, carry a few stones to the bridge that we believe can be built between single cell and human levels of cognition, by (c) suggesting ways in which individual agency is influenced by the social interactions that social agents engage in.

Notes

1 When we use the word *meaning-making*, this is synonymous with *sense-making*. *Meaning* and *sense* are also interchangeable in the present article.

2 We do have a terminological issue in this article though, which is that the word *agent* is standardly used in adaptive behavior research to refer to the artificial or simulated system whose behavior we study. In this article, we stick to this convention, and indicate which sense of *agent* we employ only in those places where we think it is likely to cause confusion.

3 Let it be clear that our view does not imply that individuals have no responsibility. Quite the contrary. The idea of “responsibility” in this kind of approach to social interaction (and its potential ethical implications) is certainly one that deserves further elaboration, but this article is not the place to do that (see also Section 4.10).

4 A question that may surface at this point (as it did for one of our reviewers, to whom we are grateful for suggesting it) is that of whether human level agency (crossing the cognitive gap) *demands* social cognition, or whether it would be possible to develop solipsistic machines with such agency? Our response would be that it might be possible to design solipsistic machines with human agency, but that the effort would be immense and unnecessary. An indication that it is possible comes from people with deficits that impact on social capacities, such as autism and Möbius syndrome (people with the latter cannot express facial gestures). Such deficits often come with concomitant developmental problems, though Cole (in press) suggests that people with Möbius can learn to compensate cognitively. The same may be true for people with autism—if they do employ something like a so-called “theory of mind,” this may be a way of compensating for what they lack in other aspects of social fluency. We propose that going the social route will be the more parsimonious option.

Acknowledgments

We would like to thank two anonymous reviewers for their helpful comments and suggestions, as well as Takashi Ikegami and Marieke Rohde, the editors of this special issue, and Ezequiel Di Paolo, for their generous help. HDJ’s part of the work was supported by the EU Marie Curie Research Training Network 035975 DISCOS—Disorders and Coherence of the Embodied Self.

References

Ashby, W. R. (1960). *Design for a brain: The origin of adaptive behavior* (2nd ed.). London, UK: Chapman and Hall.

Auvray, M., Lenay, C., & Stewart, J. (2009). Perceptual interactions in a minimalist virtual environment. *New Ideas in Psychology*, *27*(1), 32–47.

Barandiaran, X., Di Paolo, E., & M. Rohde (2009). Defining agency: individuality, normativity, asymmetry and spatio-temporality in action. *Adaptive Behavior, 17,* 367–386.

Beer, R. D. (1995). A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence*, *72*(1–2), 173–215.

Beer, R. D. (1997). The dynamics of adaptive behavior: A research program. *Robotics and Autonomous Systems*, *20*(2–4), 257–289.

Beer, R. D. (2003). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior*, *11*(4), 209–243.

Boden, M. (2006). *Mind as machine: A history of cognitive science*. Oxford: Oxford University Press.

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, *47*(1–3), 139–160.

Brooks, R. A. (1997). From earwigs to humans. *Robotics and Autonomous Systems*, *20*(2–4), 291–304.

Buck, J., & Buck, E. (1976). Synchronous fireflies. *Scientific American*, *234*, 74–85.

Cangelosi, A. (2007). Adaptive agent modeling of distributed language: investigations on the effects of cultural variation and internal action representations. *Language Sciences*, *29*(5), 633–649.

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

Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. New York: Oxford University Press.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, *58*(1), 7–19.

Cole, J. (in press). Impaired embodiment and intersubjectivity. *Phenomenology and the Cognitive Sciences*. doi: 10.1007/s11097-009-9119-5.

De Jaegher, H. (2009). Social understanding through direct perception? Yes, by interacting. *Consciousness and Cognition*, *18*(2), 535–542.

De Jaegher, H., & Di Paolo, E. A. (2007). Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences*, *6*(4), 485–507.

De Jaegher, H., & Di Paolo, E. (2008). Making sense in participation. An enactive approach to social cognition. In F. Morganti, A. Carassa, & G. Riva (Eds.), *Enacting intersubjectivity: A cognitive and social perspective to the study of interactions*. Amsterdam: IOS Press. 33–47.

de Rugy, A., Salesse, R., Oullier, O., & Temprado, J.-J. (2006). A neuro-mechanical model for interpersonal coordination. *Biological Cybernetics*, *94*, 427–443.

Di Paolo, E. A. (2000). Behavioral coordination, structural congruence and entrainment in a simulation of acoustically coupled agents. *Adaptive Behavior*, *8*, 25–46.

Di Paolo, E. A. (2003). Organismically inspired robotics: homeostatic adaptation and teleology beyond the closed sensorimotor loop. In K. Murase & T. Asakura (Eds.), *Dynamical systems approach to embodiment and sociality* (pp. 19–42). Adelaide, Australia: Advanced Knowledge International.

Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, *4*(4), 429–452.

Di Paolo, E. A. (2009). Extended Life. *Topoi*. doi: 10.1007/s11245-008-9042-3. 9–21.

Di Paolo, E. A., Rohde, M., & De Jaegher, H. (in press). Horizons for the enactive mind: Values, social interaction, and play. In J. Stewart, O. Gapenne, & E. A. Di Paolo (Eds.), *Enaction: Towards a new paradigm for cognitive science*. Cambridge, MA: MIT Press.

Di Paolo, E. A., Rohde, M., & Iizuka, H. (2008). Sensitivity to social contingency or stability of interaction? Modelling the dynamics of perceptual crossing. *New Ideas in Psychology*, *26*(2), 278–294.

Fink, P. W., Kelso, J. A. S., Jirsa, V. K., & de Guzman, G. (2000). Recruitment of degrees of freedom stabilizes coordination. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 671–692.

Fogel, A. (1993). *Developing through relationships: Origins of communication, self and culture*. London: Harvester Wheatsheaf.

Froese, T., & Di Paolo, E. A. (2008). Stability of coordination requires mutuality of interaction in a model of embodied agents. In M. Asada et al. (Eds.), *From Animals to Animats 10: Proceedings of the 10th International Conference on Simulation of Adaptive Behavior* (pp. 52–61). Berlin: Springer-Verlag.

Froese, T., Virgo, N., & Izquierdo, E. (2007). Autonomy: a review and a reappraisal. In F. Almeida e Costa et al. (Eds.), *Advances in Artificial Life: Proceedings of the 9th European Conference on Artificial Life* (pp. 455–464). Berlin: Springer-Verlag.

Froese, T., & Ziemke, T. (2009). Enactive artificial intelligence: Investigating the systemic organization of life and mind. *Artificial Intelligence*, *173*(3–4), 366–500.

Fuchs, T., & De Jaegher, H. (in press). Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenology and the Cognitive Sciences*.

Gallagher, S. (2001). The practice of mind: theory, simulation or primary interaction? *Journal of Consciousness Studies*, *8*(5–7), 83–108.

Gallagher, S. (2005). *How the body shapes the mind*. Oxford: Oxford University Press.

Gergely, G., & Watson, J. (1999). Early social-emotional development: contingency perception and the social biofeedback model. In P. Rochat (Ed.), *Early social cognition: Understanding others in the first months of life* (pp. 101–137). Hillsdale, NJ: Lawrence Erlbaum.

Gibbs, R. W. Jr. (2001). Intentions as emergent products of social interactions. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 105–122). Cambridge, MA: MIT Press.

Goffman, E. (1983). The interaction order. *American Sociological Review*, *48*, 1–17.

Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.

Grammer, K., Kruck, K. B., & Magnusson, M. S. (1998). The courtship dance: Patterns of nonverbal synchronization in opposite-sex encounters. *Journal of Nonverbal Behavior*, *22*(1), 3–29.

Granic, I. (2000). The self-organization of parent-child relations: beyond bidirectional models. In M. D. Lewis & I. Granic (Eds.), *Emotion, development, and self-organization. Dynamic systems approaches to emotional development* (pp. 267–297). Cambridge, UK: Cambridge University Press.

Harvey, I., Di Paolo, E. A., Wood, R., Quinn, M., & Tuci, E. A. (2005). Evolutionary robotics: A new scientific tool for studying cognition. *Artificial Life*, *11*(1–2), 79–98.

Hobson, R. P. (1991). Against the theory of "Theory of Mind." *British Journal of Developmental Psychology*, *9*, 33–51.

Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.

Iizuka, H., & Ikegami, T. (2004). Adaptability and diversity in simulated turn-taking behavior *Artificial Life*, *10*(4), 361–378.

Iizuka, H., & Di Paolo, E. A. (2007). Minimal agency detection of embodied agents. In F. Almeida e Costa et al. (Eds.), *Advances in Artificial Life: Proceedings of the 9th European Conference on Artificial Life* (pp. 485–494). Berlin: Springer-Verlag.

Ikegami, T., & Iizuka, H. (2007). Turn-taking interaction as a cooperative and co-creative process. *Infant Behavior & Development*, *30*(2), 278–288.

Jirsa, V. K., Fuchs, A., & Kelso, J. A. S. (1998). Connecting cortical and behavioral dynamics: bimanual coordination. *Neural Computation*, *10*(8), 2019–2045.

Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behaviour*. Cambridge, MA: MIT Press.

Kirsh, D. (1991). Today the earwig, tomorrow man? *Artificial Intelligence*, *47*(1–3), 161–184.

Lave, J., & Wenger, E. (1991). *Situated learning. Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.

Marocco, D., & Nolfi, S. (2007). Emergence of communication in embodied agents evolved for the ability to solve a collective navigation problem. *Connection Science*, *19*, 53–74.

Murray, L., & Trevarthen, C. (1985). Emotional regulations of interactions between two-month-olds and their mothers. In T. M. Field & N. A. Fox (Eds.), *Social perception in infants* (pp. 177–197). Norwood, NJ: Ablex Publishing.

Nadel, J., Carchon, I., Kervella, C., Marcelli, D., & Réserbat-Plantey, D. (1999). Expectancies for social contingency in 2-month-olds. *Developmental Science*, *2*(2), 164–173.

O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, *24*(5), 883–917.

Quinn, M. (2001). *Evolving communication without dedicated communication channels.* Paper presented at the Advances in Artificial Life: 6th European Conference on Artificial Life (ECAL 2001), Prague.

Quinn, M., & Noble, J. (2001). Modelling animal behaviour in contests: Conventions for resource allocation. In *Advances in Artificial Life: Sixth European Conference on Artificial Life (ECAL 2001)* (pp. 367–376). Berlin: Springer.

Reddy, V., & Morris, P. (2004). Participants don't need theories: Knowing minds in engagement. *Theory and Psychology*, *14*(5), 647–665.

Reed, K., Peshkin, M., Hartmann, M. J., Colgate, J. E., & Patton, J. (2005). *Kinesthetic interaction.* Paper presented at the 9th International Conference on Rehabilitation Robotics.

Reed, K., Peshkin, M., Hartmann, M. J., Grabowecky, M., Patton, J., & Vishton, P. M. (2006). Haptically linked dyads: Are two motor-control systems better than one? *Psychological Science*, *17*(5), 365–366.

Russell, J. (1996). *Agency: Its role in mental development*. Hove, UK: Taylor & Francis.

Sacks, H. (1992). *Lectures on conversation Volumes I and II*. Oxford: Blackwell.

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

Sheets-Johnstone, M. (1999). *The primacy of movement*. Amsterdam: John Benjamins.

Steels, L., & Kaplan, F. (2002). Bootstrapping grounded word semantics. In T. Briscoe (Ed.), *Linguistic evolution through language acquisition: Formal and computational models*. Cambridge, UK: Cambridge University Press.

Stewart, J. (1996). Cognition = life: Implications for higher-level cognition. *Behavioral Processes*, *35*, 311–326.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action.* London: Bradford.

Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind.* Cambridge, MA: The Belknap Press of Harvard University Press.

Varela, F. J. (1979). *Principles of biological autonomy*. New York: Elsevier North Holland.

Varela, F. J. (1995). The re-enchantment of the concrete: Some biological ingredients for a nouvelle cognitive science. In L. Steels & R. Brooks (Eds.), *The artificial life route to artificial intelligence* (pp. 11–22). Hove, UK: Lawrence Erlbaum Associates.

Varela, F. J., Maturana, H. R., & Uribe, R. (1974). Autopoiesis: The organization of living systems, its characterization and a model. *BioSystems*, *5*, 187–196.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Vygotsky, L. (1986). *Thought and language* (revised ed.). Cambridge, MA: MIT Press.

Weber, A., & Varela, F. J. (2002). Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, *1*, 97–125.

Wheeler, M. (1997). Cognition's coming home: the reunion of life and mind. In P. Husbands & I. Harvey (Eds.), *Proceedings of the 4th European Conference on Artificial Life* (pp. 10–19). Cambridge, MA: MIT Press.

Winfree, A. T. (2001). *The geometry of biological time*. London: Springer.

About the Authors

Hanne De Jaegher finished her D.Phil. in philosophy of cognitive science at the University of Sussex, UK, in 2007. She is currently postdoctoral fellow in the Marie Curie Research Training Network DISCOS (Disorders and Coherence of the Embodied Self) and works at the Department of General Psychiatry at the University of Heidelberg, Germany. She is interested in interactional approaches to social cognition (intersubjectivity) and novel methodologies for investigating these.

Tom Froese is currently completing his doctoral thesis on "Sociality and the life-mind continuity thesis: A study in evolutionary robotics" at the University of Sussex (Brighton, UK). He previously obtained an M.Eng. in computer science and cybernetics at the University of Reading (Reading, UK). His main research area is the relationship between life and mind, as expressed in living (biological) and lived (experiential) phenomena.
E-mail: t.froese@gmail.com