Phenom Cogn Sci (2012) 11:205–235
DOI 10.1007/s11097-012-9254-2

The extended body: a case study in the neurophenomenology of social interaction

Tom Froese · Thomas Fuchs

Published online: 23 March 2012

Abstract There is a growing realization in cognitive science that a theory of embodied intersubjectivity is needed to better account for social cognition. We highlight some challenges that must be addressed by attempts to interpret 'simulation theory' in terms of embodiment, and argue for an alternative approach that integrates phenomenology and dynamical systems theory in a mutually informing manner. Instead of 'simulation' we put forward the concept of the 'extended body', an enactive and phenomenological notion that emphasizes the socially mediated nature of embodiment. To illustrate the explanatory potential of this approach, we replicate an agent-based model of embodied social interaction. An analysis of the model demonstrates that the extended body can be explained in terms of mutual dynamical entanglement: inter-bodily resonance between individuals can give rise to self-sustaining interaction patterns that go beyond the behavioral capacities of isolated individuals by modulating their intra-bodily conditions of behavior generation.

Keywords Enaction · Embodied intersubjectivity · Dynamical systems theory · Social cognition

Introduction

We set the stage for our approach to sociality with a critical review of the mainstream literature in the sciences of the mind. We argue that despite some prominent critiques

T. Froese (✉)
Ikegami Laboratory, The University of Tokyo,
Room 225b, Building 16, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan
e-mail: t.froese@gmail.com

T. Fuchs
Klinik für Allgemeine Psychiatrie, Zentrum für Psychosoziale Medizin,
Universitätsklinikum Heidelberg, Voß-Str. 4, 69115 Heidelberg, Germany
e-mail: thomas.fuchs@med.uni-heidelberg.de

of traditional cognitivist accounts and attempts to take embodiment, social interaction and intersubjectivity into consideration, the core of most of this research remains securely tied to the classical cognitivist 'theory of mind' framework. In particular, we single out theories centered on the discovery of 'mirror' neurons and show that they continue to accept an overly restrictive conception of the conscious mind, which is characterized by what has been called "The Myth of the Inner; The Myth of the Hidden; and The Myth of the Single" (Torrance 2009: 112). We argue that in spite of the increased attention given to the role of the body and the neurophysiological foundations of sociality, for instance in the case of 'embodied simulation' (Gallese 2005), the field of social cognition generally remains attached to the cognitivist framework. This approach is characterized by an insistence on homuncular explanations, especially in terms of sub-personal mental representations, theorizing, and/or pretense. Moreover, because of the cognitivist assumption that the mind is hidden in the head, the field of social cognition has difficulties in moving beyond the confines of the brain of an isolated individual.

After uncovering this lingering cognitivist core in the literature, even in the case of some embodied accounts, we turn toward what we consider its most promising aspects, namely a greater acceptance of phenomenological insights and a growing scientific interest in the role of self-other relations for the mind. In the rest of this paper we demonstrate how these two trends can be fruitfully combined in a way that better avoids the problems associated with cognitivism.

The standard approach

Since the days of the 'cognitive revolution' in the 1960s, which gave birth to cognitive science with the computational theory of mind, the topic of sociality has been treated under the general rubric of 'social cognition' (Fiske and Taylor 1991). The general aim of this field has been to uncover how cognitive agents are able to think about themselves and others (e.g. Fodor 1992; Baron-Cohen et al. 1985; Premack and Woodruff 1978). In particular, the focus has been on finding computational mechanisms that are constitutive of social information processing. The most famous example is the search for a 'theory of mind mechanism' that enables an observer to 'mentalize' or 'mindread', i.e. to establish the presence of hidden mental representations of conspecifics. While there has been an ongoing debate about the precise nature of this mechanism, such as whether it is best described as akin to a scientific 'theory' or rather as a mental 'simulation' routine (Stich and Nichols 1991; Carruthers and Smith 1996), the general framework of the discussion has mostly remained true to its origins in classical cognitivism. Briefly, the debate centers on whether our understanding of others is primarily based on theorizing, i.e. on making hypotheses about the states of another's mind and testing them in a process that is akin to scientific practice, or on simulating, i.e. evaluating the states of another's mind by putting oneself into their situation in an imagined transferal. By explaining social understanding primarily in terms of theorizing and/or imagining, these theories are committed to a view of sociality as based on higher-level abstract cognition.

However, major changes to this received view are currently underway, as evidenced by this and other recent collections on alternative approaches to sociality in the sciences of the mind.[1] Beginning in the early 90s, these changes are driven by a confluence of factors, most notably by the increasing acceptance of embodied, enactive, extended, dynamical, externalist and phenomenological accounts of mind.[2] Regarding alternative approaches to sociality more specifically, so far the most prominent driving force has been the empirical discovery of the so-called 'mirror' neurons (Rizzolatti et al. 1996), i.e. neurons that tend to be active in cases of both action execution and action perception.[3] This discovery has reinvigorated the old cognitivist theory of mind controversy between 'theory theory' versus 'simulation theory', with the consensus now centering on the latter theory. Moreover, the terms of the debate have begun to noticeably shift from the concepts of classical cognitivism toward a more embodied conception of social interaction. Thus, classical explanations framed in terms of propositional attitudes such as 'intentions', 'beliefs' and 'desires' are no longer fashionable (Gallese 2007: 659), especially if they are intended to describe the operation of sub-personal mechanisms. Instead, it is popular to cite the 'mirror' neuron discovery in support of the role of embodiment for social understanding, for example in terms of "resonance behaviors" (Rizzolatti et al. 1999), "motor resonance" (Agnew et al. 2007), "motor cognition" (Jackson and Decety 2004), "motor embodiment" (Keysers and Fadiga 2008) and "embodied simulation" (Gallese 2005).

It has not gone unnoticed, however, that this change of terminology hides a continuing affinity with the basic assumptions of classical cognitivism. For instance, notwithstanding the increased appeal to the notion of embodiment, it is not clear to what extent the body plays a constitutive role in these more implicit, sub-personal versions of simulation theory, except perhaps as traditional mental representations realized in "bodily formats" (Goldman and de Vignemont 2009). In other words, as implied by the popular metaphor of the 'mirror', the function of these neurons is discussed as if their activation was *re-presenting* external states of affairs within the brain. This concern with mental representation is, of course, familiar classical cognitivist territory, even if the motor system is now assumed to participate in those representational networks. Moreover, this neural representation is supposedly couched in *pretense*, since otherwise it would be a realization and not a 'simulation', and it is said that these neurons are actually firing *as if* they were in fact involved in executing the action. However, given that pretense behavior is a sophisticated social skill, it is simply absurd to assume that a neuron or even a neural assembly is capable of pretense, i.e. of pretending to be discharging (Gallagher 2007). It does not make sense to explain our socio-cognitive skills by attributing those very same skills to our brains. Thus it seems that, despite the terminological changes, the notion of 'embodied simulation' is still haunted by the cognitivist homunculus.

In order to overcome this lingering cognitivism, Gallagher (2008) suggests that a better way of interpreting the activity of neurons that are active during action execution as well as action observation is in terms of the phenomenological notion

[1] The interested reader is referred to, for example, Thompson 2001; Stawarska 2006; Zlatev et al. 2008; Morganti et al. 2008; Di Paolo 2009; Hutto and Ratcliffe 2010.

[2] See, e.g., Varela et al. 1991; Thelen and Smith 1994; Port and van Gelder 1995; Clark 1997; Noë 2004; Wheeler 2005; Gallagher 2005; Thompson 2007; Stewart et al. 2010.

[3] See, e.g., Rizzolatti et al. (1999, 2001); Gallese et al. (2004);

of ‘direct perception’, combined with an ‘enactive’ or sensorimotor approach to perception. Phenomenology has long questioned the assumption that the minds of others are completely hidden from our perception and locked away in an inner mental realm (Zahavi 2001). In addition, some phenomenologists, as well as the enactive approach, have insisted on the interdependence of action and perception (Varela et al. 1991). To put it differently, if social understanding implies the perception of the other, and if there is no clear separation between action and perception, then we should expect to find the involvement of ‘motor’ neurons during an agent’s perception of the activity of other agents. The evidence of ‘mirror’ neurons can therefore also be used to support the idea that perception of others is a process of enaction, rather than simulation. This interpretation neatly exorcises the cognitivist homunculus by appealing to phenomenological evidence, and dispenses with the problematic notions of neuronal pretense and ‘mirroring’.

There is another tension in the ‘mirror’ neuron literature that has to do with the problem of what defines the ‘social’ of social cognition. One of the defining premises of classical research is ‘methodological individualism’, a sociological concept that has been adapted by Boden (2006) to describe an assumption of mainstream cognitive science according to which the individual cognitive agent is the proper unit of analysis of mind and behavior, including the case of social interaction and other collective phenomena.[4] On this view, it is possible to get a complete understanding of the mechanisms underlying an agent’s social behavior by focusing on that agent alone.

However, this solipsistic doctrine is being increasingly challenged from a variety of alternative approaches in the sciences of the mind,[5] and the ‘mirror’ neuron literature at first sight appears to be part of this trend. In addition to frequent references made to “resonance” between self and other (Rizzolatti et al. 1999), it is argued that by means of these neurons “a bridge is created between others and ourselves” (Gallese et al. 2004: 400), that they “sustain shared representations between self and other” (Jackson and Decety 2004: 259), and thus ultimately contribute to “intentional attunement” (Gallese 2007). But as radically non-classical as some of these proposals may sound, a closer look reveals that they are still bound by the cognitivist assumption of methodological individualism.

For example, in the case of ‘embodied simulation’ the mind of the other agent is internally simulated in terms of the self. And even in interpretations that are self-other ‘neutral’, for instance de Vignemont’s (2004) proposal of “co-consciousness”, we find tensions created by cognitivist hangovers. Such a neutral approach initially shows potential, namely by hypothesizing that “we do not need to start from the self in order to go to the other; rather we can start from what is common between oneself and the other” (ibid.: 108), but it ends in confusion. It is worth quoting a passage from this proposal at length in order to highlight the difficulties that are encountered when

[4] The phrase ‘methodological individualism’ is often used differently in the philosophy of mind, where it follows usage that was first introduced by Fodor. We use the phrase following Boden and the tradition of Weber, Hayek, and Popper. See the discussion by Heath (2011).

[5] See, e.g., Hutchins (1995); Ratcliffe (2007); De Jaegher and Di Paolo (2007); Di Paolo et al. (2008); Reddy (2008); Fuchs (2008); Gallagher (2008); Hutto (2008); Auvray et al. (2009); Fuchs and De Jaegher (2009); Starwaska (2009); Froese and Di Paolo (2011a); Stuart (2011); Torrance and Froese (2011).

trying to approach the notion of embodied intersubjectivity by starting from within the cognitivist framework:

> Gopnik suggests that we have the same representations of mental states for ourselves and for the others from a third-person point of view. However, this does not sound plausible. Indeed, we may sometimes use the same abstract representations for ourselves and for the others, but we cannot eliminate the subjective perspective that accompanies self-knowledge, and more particularly in the case of emotions: there is something it is like to feel happiness from the point of view of the subject. What I am suggesting here is that rather than reducing self-knowledge to a specific kind of theoretical knowledge of other minds, as Gopnik tries to do, we should bring knowledge of others closer to self-knowledge: we share common representations of ourselves and the others from a first-person point of view. Mirror neurons do not represent actions from an abstract perspective, but rather from the intersubjective point of view. Thus, from a first-person perspective we are able to represent actions preformed by someone else as if they were our own and there is inference from you to the other involved here. (de Vignemont 2004: 109–110)

De Vignemont makes an attempt to mitigate cognitivism's traditional explanatory focus on abstract theoretical inference by appealing to phenomenology and neuroscience. But it is difficult to make sense of this position, especially because of the lingering cognitivist tendency of indiscriminately applying the concept of 'mental representation' to personal, interpersonal and subpersonal levels of description. To insist that a human subject represents actions from the first-person point of view, while simultaneously saying that a mirror neuron represents actions from the intersubjective point of view, only has the effect of leaving it mysterious as to what could be meant by the terms 'representation' and 'point of view'. If both human subjects and mirror neurons represent these states of affairs from a point of view, is there no mental difference between a person and their neuron? And even if we restrict the notion of representation to the personal level alone, there is still phenomenological confusion: What does my feeling of happiness 'represent' to my point of view? In general, the hard problem of consciousness will not simply go away by re-describing subjective experience in terms of inner mental representations, and then ascribing these mental representations to neurons. Phenomenology and neuroscience can indeed be brought into a meaningful relationship with each other, but only if we take into account that we are dealing with two distinct domains of evidence. These domains are certainly not independent from each other, but they are non-trivially related such that they maintain their own unique properties.

Leaving aside the problematic use of established phenomenological concepts without a clear personal/subpersonal distinction, de Vignemont's basic proposal seems to be to replace the traditional role of abstract theoretical inference in social cognition with first-person experience, which would align this account with phenomenology. However, there is an additional cognitivist stumbling block: although knowledge of others is related to the existence of an intersubjective point of view, the notion of having a 'point of view' is accounted for only in terms of self-knowledge so that, ultimately, knowledge of others is again reduced to the cognitive self. Social interaction with others plays no essential role in this account of intersubjectivity, and the neutral account of ‘co-consciousness’ turns out to be another version of methodological individualism after all.

Importantly, the problem that no consideration is given to the potential roles of social interaction can be generalized to most theories of sociality whose foundations rest on ‘mirror’ neurons [MNs]. Jacob and Jeannerod (2005: 23) do well to emphasize that “MNs were first discovered in the context of motor and perceptual tasks that had a very weak social content, if any.” More precisely, the firing activity of these neurons is in fact entirely independent of whether or not any social interaction is taking place.

> When MNs fire in the brain of a monkey watching another grasp a fruit, the discharge is a weakly social process: the two monkeys are not involved in any kind of non-verbal intentional communication. (Jacob and Jeannerod 2005: 21)

The starting point of the ‘mirror’ neuron literature is therefore the same as that of social cognition in classical cognitivism: a passive observer is presented with an independent external stimulus, which in this case happens to be another agent, and the aim is to give an internal mechanism that explains the observer’s behavior. In other words, since it is assumed that the other agent’s mind is completely hidden from perception, thus requiring cognitive access of some kind, and since the observer’s cognitive mechanism is assumed to be internal to its brain, we end up with some version of a neural module account of an individual’s ‘mindreading’ abilities. Despite initial appearances to the contrary, the recent ‘mirror’ neuron theories of sociality remain fundamentally based on cognitivist assumptions about the mind. They implicitly remain committed to the myth of the hidden, the myth of the inner, and the myth of the single (Torrance 2009).

An alternative approach

We have uncovered some significant tensions in current mainstream approaches to social cognition. Classical cognitivist ‘theory theory’ has become unfashionable, but a distinct alternative remains as yet to be established. To be sure, the notion of ‘mirror’ neurons has raised the prospect of developing a new theory of sociality, one that takes the role of embodiment and social interaction into account. However, a lingering commitment to the philosophical assumptions and explanatory principles of classical cognitivism impedes further progress in this direction.

One important point is that cognitive scientists must pay closer attention to the mental phenomena that they are aiming to explain. Folk psychology is not a reliable source of insights (Ratcliffe 2007). The essential features of our experience of social phenomena must be systematically determined by careful phenomenological investigation, thus avoiding explanations of social cognition that may look theoretically elegant, but which remain detached from reality. However, although phenomenology’s traditional emphasis on descriptive accounts of lived experience has given rise to a rich source of concepts, their prose-based format poses a challenge to scientific integration. And how can phenomenology be used to inform an alternative account of what happens at the subpersonal level? A key insight into resolving this challenge is that one of the most fundamental structures of consciousness is temporality, and that the form of this time-consciousness can also be described in terms of the mathematics of dynamical systems theory. This insight has been put to good use in the science of consciousness (e.g. Varela 1999; van Gelder 1999). In addition, the scientific reformulation of phenomenological descriptions of experiential structures in precise mathematical terms, which are amenable to further formal analysis, may become a starting point for new phenomenological clarifications as well. At the very least, they offer an opportunity to return to the phenomena themselves in order to get more precise descriptions.

Moreover, dynamical systems theory, which is the standard mathematical framework of the natural sciences, has also been successfully applied in cognitive science (Beer 2000; Port and van Gelder 1995), thus forming a convenient interdisciplinary bridge from which to explore the mind-body problem. Varela (1996) has pioneered the use of dynamical systems theory in combination with phenomenology and cognitive science with his neurophenomenological research program. Neurophemenology was conceived as a methodological renewal of the cognitive sciences, and it has been successfully applied in different areas.[6] Thus, instead of starting with abstract ideas derived from computer science and the computational theory of mind, and then applying these concepts to explain experiential and empirical observations, neurophenomenology proceeds the other way around: it begins with concrete phenomenological and scientific insights, and then describes their interlinking and shared temporal structures in terms of dynamical systems theory. The aim is to establish a mutual circulation of insights (Fig. 1).

The mathematics of dynamical systems provides a formal way of *describing* temporal structures, but in itself it does not amount to a scientific *theory*. One promising approach toward a more adequate theory of life, mind and sociality is to replace the traditional metaphor of the cognitive sciences, i.e. computation, with that of 'enaction' (Stewart et al. 2010). One of the basic principles of this approach, namely that an experiential world is brought forth or 'enacted' by an embodied subject, is currently being refined so as to more explicitly acknowledge the constitutive role played by other agents (Froese and Di Paolo 2011a). The guiding hypothesis is that through our mutual interactions with others our living and lived bodies become inextricably intertwined in a dynamical whole, thus forming an 'extended body' by which we enact and encounter the world together (Froese 2011). In order to study this active role of the interaction process itself, new experimental methodologies and conceptual tools of analysis are being developed (De Jaegher et al. 2010). In what follows we contribute to this ongoing endeavor (1) by clarifying some essential aspects of the phenomenology of the extended body, and (2) by working toward a formalization of these aspects with the help of a dynamical systems model of embodied social interaction. The modeling results will demonstrate that we can make the notion of embodied intersubjectivity intelligible in a non-reductive scientific manner, but without recourse to the concepts of inner representation or simulation.

[6] For a more detailed discussion of this methodology: Varela et al. (1991); Gallagher (1997); Varela (1997); Roy et al. (1999); Lutz and Thompson (2003); Di Paolo et al. (2010); and Froese et al. (2011). Applications of this methodology can be found in: Varela (1999); Lutz (2002); and Petitmengin et al. (2007).

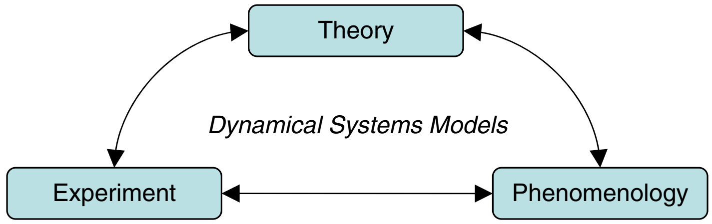


Fig. 1 The three pillars of an alternative framework for cognitive science: theory, experiment, and phenomenology are integrated into one coherent methodology by relations of mutual enlightenment and constraint. In some cases they can be formally brought together by means of dynamical systems models

Phenomenology of the extended body

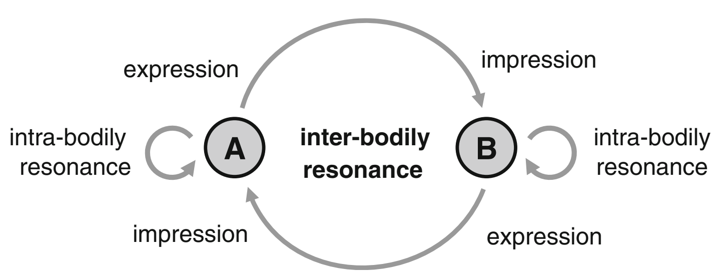
We begin our case study by clarifying the phenomenology of embodied intersubjectivity, in particular of what we call the 'extended body'. Although inter-bodily interactions are typically experienced in a pre-reflective manner that hides them from direct observation and rational scrutiny, they still leave a noticeable mark on our moods and feelings as well as on our lasting dispositions. We will have a closer look at two facets of the extended body, namely (a) inter-bodily resonance, and (b) inter-bodily memory.

a. Inter-bodily resonance

First, the phenomenon of the extended body may be described as a form of pre-reflective face-to-face interaction, or embodied communication, which is taking place on the basis of an implicit *inter-bodily resonance* (Fuchs 1996) as illustrated in Fig. 2.

Let us assume that A is a person whose emotion, e.g. anger, manifests itself in typical bodily (facial, gestural, interoceptive, adrenergic, circulatory, etc.) changes. His pre-reflectively experienced 'lived' body thus functions as a felt 'resonance board' for the emotion: A feels the anger *as* the tension in his face, *as* the sharpness of his voice, the arousal in his body etc. These proprio- and interoceptive bodily feelings may be termed *intra-bodily resonance.* However, this resonance is an *expression* of the emotion at the same time, that means, the anger becomes visible and is perceived as such by A's partner B. But what is more, the expression will also produce an *impression*, namely by triggering corresponding or complementary bodily feelings in B. Thus, A's sinister gaze, the sharpness of his voice or expansive bodily movements might induce in B an unpleasant tension or even a jerk, a tendency to withdraw, etc. Thus, B not only sees the anger immediately in A's face and gesture, but also senses it with his own body, through his own intra-bodily resonance.

Fig. 2 An illustration of the extended body (figure adapted from Fuchs 1996)



However, it does not stay like this, for the impression and bodily reaction caused in B in turn becomes an expression for A; it will immediately affect his bodily reaction, change his expression, however slightly, and so forth. This creates a circular interplay of expressions and reactions running in split seconds and constantly modifying each partner's bodily state, in a process that becomes highly autonomous and is not directly controlled by the partners. They have become parts of a dynamic sensori-motor and inter-affective system that connects both bodies by reciprocal movements and reactions, that means, in *inter-bodily resonance*. Of course, the signals and reactions involved proceed far too quickly to stand out discretely and become conscious as such. Instead, both partners will experience a specific feeling of being connected with the other in a dynamic way that may be termed "mutual incorporation" (Fuchs and De Jaegher 2009). Each lived body reaches out, as it were, to be extended by the other. This is accompanied by a particular, holistic impression of the interaction partner and a feeling for the overall atmosphere of the shared situation.

No mental representation or simulation is necessary for this process, at least not at the experiential level. We certainly do not simulate, for example, another's angry gaze or voice, even less his anger, but rather feel tense, threatened or invaded by his expressive bodily behaviour. There is no strict or dualistic separation between the inner and the outer at all, as if a hidden mental state in A produced certain external signs that B would have to decipher. A's anger may not be separated from its bodily resonance and expression; and similarly, B does not perceive A's body as a mere object, but as a living, animate and expressive body that he feels more or less connected with. In other words, the *lived body's* impression on the one side (A) becomes a *living body's* expression for the other (B), and vice versa: the impression produced in B's lived body becomes a living body's expression for A. In both partners, their own intra-bodily resonance is fed into the perception of the other. Bodily sensations, tensions, action tendencies, etc. that arise in the interaction do not serve as a separate simulation of the other person, but are part and parcel of their mutual perception. It is in this sense that we can refer to the experience of the other in terms of 'embodied' perception, which, through the interaction process, is at the same time an 'embodied' communication.

b. Inter-bodily memory

This account of inter-bodily resonance may be complemented by the historical dimension. In embodied agents, the history of interactions continuously changes their dispositions. From early childhood on, patterns of interaction are sedimented in the infant's implicit or bodily memory, resulting in what may be called inter-bodily or 'intercorporeal memory' (Fuchs 2008; Fuchs and De Jaegher 2009; Fuchs 2012). This means a pre-reflective, practical knowledge of how to interact with others—e.g. how to share pleasure, elicit attention, avoid rejection, re-establish contact, etc. It may also be termed "implicit relational knowledge" (Lyons-Ruth et al. 1998). As such, intercorporeal memory enables the basic formation of dyadic and more generally intersubjective patterns of interaction.

Even the earliest experiences of how infants are held, comforted, guided and reacted to by their care-givers are imprinted in their body memory, hence also displayed in their later actions and their habitus, i.e. their entire set of learned dispositions. Moreover, the infant acquires specific interactive schemes ('schemes of being-with', Stern, [1985] 1998) and bodily micro-practices (Downing 2004) that are needed for a growing range of interactions. Inter-bodily memory is a temporally organised, 'musical' ability to engage in the typical rhythms, dynamics and affects that are present in the interaction with others, already beginning in infancy (Trevarthen 2005). When it is actualized, it shapes the present relationship as a procedural field that encompasses and connects both partners.

Thus, as a result of learning processes, which are in principle comparable to the acquiring of motor skills, social agents shape and enact their relationships according to the patterns they have extracted from earlier and earliest experience. On the other hand, each particular interaction also acquires its own history, thus pre-figuring and constraining future interactions between the respective partners. When encountering each other again, both interact and behave in a way they would not outside of the relationship. Hence, we may say that there is an inter-bodily memory of the interactive process itself.

Summary

Summarizing these analyses, there appears to be a convergence between the alternative theories of social cognition and the phenomenology of embodied intersubjectivity, namely that our lived and living bodies can become extended such that they are essentially intertwined with those of others in a way that prevents any conceptual or ontological reduction to the isolated individual bodies. This applies both to current interactions and to the history of interactive patterns.

Let us take two exemplary quotes by Merleau-Ponty:

> "The communication or comprehension of gestures comes about through the reciprocity of my intentions and the gestures of others, of my gestures and the intentions discernible in the conduct of other people. *It is as if the other person's intentions inhabited my body and mine his*" (Merleau-Ponty [1945] 1962: 185; emphasis added).
>
> "In sum, our perceptions [of other agents] arouse in us *a reorganization of motor conduct, without our already having learned the gestures in question*" (Merleau-Ponty [1960] 1964: 145; emphasis added).

On this view, how we understand others is not the result of theoretical deduction or mental simulation, as mainstream cognitive science would have it. The possibility of our social understanding is grounded in a pre-reflectively lived inter-bodily reciprocity that creates a "mixture of myself and the other" (Merleau-Ponty [1960] 1964: 155). That is, the potential development of reflective knowledge about the minds of others (what the cognitivists call a 'theory of mind') depends on a more fundamental prior unification of bodies into a kind of extended body, but a unification, which nevertheless respects the differences between individual embodiments (i.e. a 'chiasmic' relationship, according to Merleau-Ponty).

The theoretical upshot of this phenomenological account is that, rather than conceiving of minds as disembodied and independent from each other, we instead treat them as fundamentally embodied and co-dependent. The enactive approach to cognitive science agrees with phenomenology on this point (Fuchs and De Jaegher 2009). But the crucial question is: can this theoretical co-dependency also be turned into a workable scientific hypothesis? Can methodological individualism be replaced by what we could call methodological 'inter-enaction' (Torrance and Froese 2011)? In what follows we will show that the extended body can in fact be experimentally studied and analytically modeled.

A model of the extended body

The focus of neurophenomenology has traditionally been on actual neuronal dynamics, and this is where one of its biggest attractions lies. Nevertheless, its guiding motivation of making phenomenology scientifically tractable in terms of dynamical systems theory can be conceived more broadly. Here we follow Froese and Gallagher (2010) in extending the existing methodology in order to include models of agent-environment interactions that are implemented on a computer in terms of nonlinear dynamical system models.

Previous work in modeling social interaction

Computer models can help us to study the dynamics of complex systems. In cognitive science, for example, they can be used to challenge and probe our intuitions of the necessary conditions of cognitive behavior (Beer 2003). These models, which are often generated automatically by evolutionary optimization so as to bracket our own design assumptions as best as possible, can serve as a useful "technological supplementation" of eidetic phenomenological methodology (Froese and Gallagher 2010).

Of course, we are not claiming that subjective experience can be reduced to these dynamical aspects, or that the modeled 'agents' themselves are subjects with experiences. The models are better conceived as a type of thought experiment that enables us to probe our intuitions and to refine our concepts (Di Paolo et al. 2000). What we are looking for are potential third-person descriptions of the kinds of conditions, which could enable the phenomenological characteristics. In addition, the mathematical insights we gain by analyzing the dynamics of such minimal systems can be helpful when it comes to understanding empirical data on actual human interactions (Di Paolo et al. 2008).

Within the context of this field of modeling research there has been a growing interest in investigating the dynamics of social interaction or 'primary intersubjectivity' (Trevarthen 1979). Some models are inspired by psychological experiments.[7] One starting point for such an endeavor is Murray and Trevarthen's (1985) classic 'double TV monitor' experiment, in which 2-month-old infants were animated by their mothers to engage in behavioral coordination via a live video link. However, when the live video of the mother was replaced with a video playback of her previously recorded actions, the infants became distressed or removed from the interaction. These results, and those of follow-up studies (e.g. Nadel et al. 1999), indicate that 2-month-old infants are sensitive to so-called 'social contingency', i.e.

[7] See, e.g., Iizuka and Di Paolo (2007); Ikegami and Iizuka (2007); Di Paolo et al. (2008); Froese and Di Paolo (2008, 2010, 2011b).

each other's mutual responsiveness during an ongoing interaction, and that this sensitivity plays a fundamental role in the unfolding of the interaction process.

Here we have a psychological study that could support the notion of the extended body that we analyzed in phenomenological terms. However, traditional explanations of these infants' sensitivity to social contingency have largely focused on individual and internal cognitive abilities. For example, Gergely and Watson (1999) have postulated the presence of an innate cognitive module which enables the detection of social contingency. These theories are defined by methodological individualism because they assume that specialized cognitive capacities exist inside of the infant in order to explain the empirical results. But it may also be possible to propose an alternative explanation that includes consideration of the interaction process as such. In other words, a more parsimonious explanation may be based on the extended body and on intercorporeal resonance, rather than on the presence of innate cognitive capacities.

In order to address this issue we can consider an illustrative computer model by Iizuka and Di Paolo (2007), who wanted to find out whether simpler solutions for the detection of social contingency could also emerge from the dynamics of the interaction process itself. In their computer simulation two 'embodied agents' were evolved to acquire the capacity to behaviorally discriminate between a 'live' (two-way interaction) condition and a 'playback' (one-way interaction) condition. A detailed analysis of one of the best solutions found by the evolutionary algorithm demonstrates that this capacity cannot be reduced to an isolated individual agent, because the dynamics of the interaction process itself play an essential role in bringing about this behavior. It is important to emphasize that the fitness evaluation used by the evolutionary algorithm did not directly specify that the solution to the task had to involve a distributed interaction process; it is just that this spontaneously happened to be one of the best strategies under the given conditions.

These results have subsequently been further generalized by Froese and Di Paolo (2008), and the essential role of the interaction process for individual behavior is supported by other simulation studies of social interactions.[8] Although a complete understanding of such computer models is often difficult, in all of these cases it is likely or actually demonstrated that the neural networks of these simulated agents do *not* utilize any kind of distinct internal representationalist mechanisms, such as social contingency detection 'modules'. Indeed, contrary to what would be expected by methodological individualism, their performance crucially depends on the dynamical properties of their mutual coupling (such as its relative stability). The models demonstrate that, at least in principle, it is not necessary to postulate special internal cognitive mechanisms to account for an agent's sensitivity to social contingency, thus extending the range of testable explanations that can also be applied to actual psychological cases. The basic take-home message is: *when we consider two or more interacting nonlinear dynamical systems (*e.g. *two agents) as a single, integrated system (*e.g. *an interaction process), we can observe novel systemic properties that cannot be reduced to the properties of the isolated components.* This claim is valid for simulated and real systems.

[8] See, e.g., Ikegami and Iizuka (2007); Di Paolo et al. (2008); Froese and Di Paolo (2010, 2011b).

This basic dynamical systems theoretic insight is one way in which the emergence of novel phenomena (e.g. behavior, cognition) from the coupling between brain, body, and world (which can include other agents) can be made intelligible. In its current form, however, it is too general to be directly useful for a better understanding of the extended body. In other words, there will indeed be novel properties that emerge out of the process of social interaction between agents, but it still remains to be shown that at least some of these properties are relevant for explaining the subsequent unfolding of the interaction process. For a strongly interactive account, there must also be mutual co-determination between the local and the global conditions. What we need to analyze in more detail is the particular situation in which embodied agents coordinate their behaviors in order to establish an interaction process, such that the process itself spontaneously modulates the individuals' conditions under which these behaviors are being generated. Only when the properties of the collective dynamics feed back into the level of behavior generation of the individual agents, and thereby form an operationally closed system with autonomous dynamics of its own (Froese and Di Paolo 2011a), do we speak of an extended body.

A model of the extended body

A mathematical dynamical system is a type of abstract model that can be used to describe the changes of a concrete phenomenon in terms of an appropriate set of variables whose states change as a function of time. There is a trade-off between making a highly detailed model that closely corresponds to the real phenomenon as much as possible, and making a mathematically tractable model that one can understand as best as possible. Since we are interested in making a general point about a general class of phenomena, we will go to the latter extreme and create a minimal model of only the most essential features. A high level of abstraction ensures that the model is applicable to a range of instances, at least at some level of description. In general, any real phenomenon can be described by *many* different mathematical dynamical systems, and each of these systems is a model of a different aspect of the various changes exhibited by the phenomenon (Giunti 1995).

The essential features of the 'extended body' have already been illustrated above. First, there must be at least two embodied agents (we refer to them as agents A and B) in a shared environment, each with a capacity for producing expressions and receiving impressions. Second, the body of each agent must consist of distinguishable parts, which can mutually influence each other's activity. This a necessary condition for the possibility of intra-bodily resonance. Third, the agents must be able to interactively coordinate their behaviors such that a recursive interaction process can be established. This is a necessary condition for the possibility of inter-bodily resonance.

We chose to formally analyze this set of minimal requirements by re-implementing a computer model by Froese and Di Paolo (2008).[9] The environment is represented as an open-ended one-dimensional line, along which the agents can move continuously.

[9] We made a new software implementation of the model by Froese and Di Paolo (2008) for the purposes of the current analysis. The parameters of the newly modeled system are identical with the old ones, except that we chose to use a more fine-grained integration method. Instead of using Euler's method with step-size 0.1 we chose a Fourth-Order Runge-Kutta with step-size 0.01. However, as should be expected, this change did not qualitatively affect the results of the model.

Each agent's body takes up 40 arbitrary units of space, an on/off contact sensor is located in the middle of the body, and the bodies are oriented so as to face each other (Fig. 3).

The temporal structure of the bodily activity of each agent is described by a mathematical dynamical system. In this case it is modeled using a continuous-time recurrent neural network (CTRNN), consisting of three fully interconnected neurons (N=3) with self-connections (Beer 2003). Each neuron is linked to a single input parameter, namely the binary state of the contact sensor. We assume that the general structure of the CTRNN is identical for both agents, although the activation states of each CTRNN will of course be different for each agent depending on their initial conditions and interaction history.

Although the components of this type of dynamical system are loosely modeled on the dynamics of real neurons (see Appendix 1), the equations are abstract enough so as to be applicable to many other kinds of systems. Thus, although we follow the tradition in evolutionary robotics of referring to these components as 'neurons', it should be kept in mind that each component could just as easily describe the overall activity of a whole system of neurons, or the state of the brain and various internal organs. The fact that we describe internal activity in terms of 'neurons' should therefore not be misunderstood as implying a brain-centered approach. Behavior is a relational phenomenon emerging from the interactions between the embodied agent as a whole and its environment. Accordingly, a CTRNN models the internal dynamics that are produced by the agent's entire body.

Given this general model setup, Froese and Di Paolo (2008) automatically configured the structure of the agents' dynamical system by using an evolutionary algorithm, which is commonly employed in the fields of artificial life and evolutionary robotics (Beer 1995, 2003; Harvey et al. 2005). The aim of this optimization was to get the model agents to exhibit behavioral coordination, which in this case required the establishment and maintenance of a common direction of movement. Importantly, the agents are not provided with any explicit communication channels over which they can transmit encoded messages, so the negotiation of a shared direction of movement has to take place within the context of the interaction itself. Following the work of Quinn et al. (2003), in this way the optimization of the interaction process has to give rise to implicit bodily communication. For more technical details about the computer implementation of the model, please refer to Appendix 1.

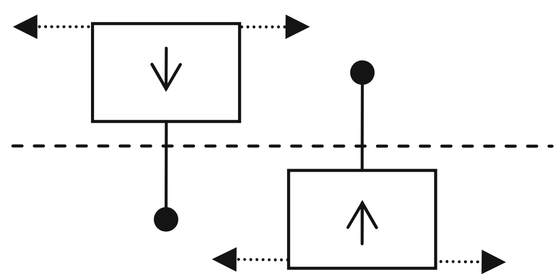


Fig. 3 An illustration of the model's basic setup showing the two agents, Agent 'A' (*facing upwards*) and Agent 'B' (*facing downwards*). The agents have 40-unit wide bodies and are placed in an open-ended one-dimensional environment. They are capable of moving left and right in a continuous manner. Their single contact sensor is located in the middle of their body

Behavioral analysis of the model agents

Since Froese and Di Paolo's (2008) computer model provides a fitting proof of concept for the notion of the extended body, we re-implemented it and conducted a detailed analysis of its properties. The agents are able to locate each other, to converge on a common direction of movement, and to sustain their interaction while moving in that direction. Time series of two trial runs, which are illustrative of the two possible directions of coordinated movement (leftwards and rightwards), with agent A and agent B starting from position 15 and 0, respectively, are shown in Fig. 4. All neurons receive input from the contact sensor. Neuron 2 and 3 control leftward and rightward movement, respectively, while neuron 1 is a general intermediate neuron.

The time series show that the agents manage to coordinate their movements by bumping their bodies into each other at more or less regular intervals. There is a significant amount of interaction between the agents, as indicated by the frequent

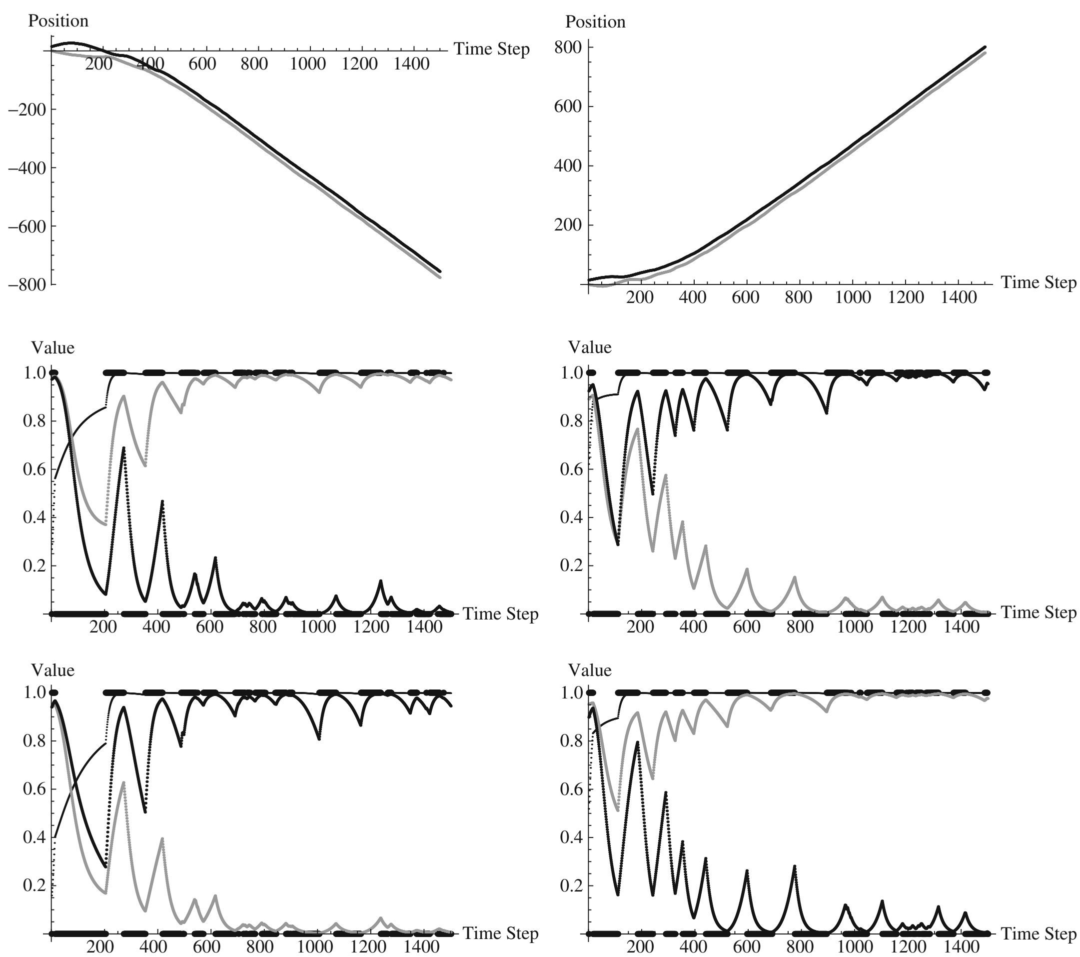


Fig. 4 Illustration of time series of the first 1500 time steps of two trials, which are representative of coordinated movement in a leftward (*left column*) and rightward (*right column*) direction, respectively. *Top row:* Positions of agent A (*black line*) and agent B (*gray line*). *Middle row:* Agent A's values of contact sensor (*thick black line*, at either 0 or 1), neuron 1 output (*thin black line*), neuron 2 output (*gray line*), and neuron 3 output (*black line*). *Bottom row:* Agent B's values of contact sensor (*thin black line*), neuron 1 output (*thick black line*), neuron 2 output (*gray line*), and neuron 3 output (*black line*)

activations of the contact sensor. Note that the longer initial sweeps of movement are useful when the agents need to locate each other from starting positions that are further apart.

In order to test the robustness of the interaction process, we also measured the ability of the agents to cope with a significant increase in the environmental perturbations by increasing the extent of the 'motor noise' (i.e. size of the random displacements that are applied to their movements). However, as shown by the left-hand graph of Fig. 5, even when the noise displacement variance was significantly increased (up to over 40 times the original magnitude used during evolution), the agents are still able to coordinate their movements effectively most of the time. The fact that the agents occasionally fail to effectively coordinate their movements in some trials shows that this is a non-trivial task.

In order to demonstrate that the success of the interaction is in fact achieved with the help of mutual responsiveness, we applied a variation of Murray and Trevarthen's 'double TV monitor' paradigm. We ran a trial as normal, but at the same time we recorded the movements of one of the agents (agent B). We then restarted the trial with the same initial conditions, but replaced the movements of agent B with a playback of its movements of the trial that was previously recorded.[10] In other words, agent A and agent B start from the same initial positions, agent A has the same initial internal activations, and agent B's movements are exactly the same as before, but agent B's movements are no longer dependent (or 'socially contingent') on the unfolding interaction with agent A. It turns out that when agent A, the 'live' agent, is confronted with this playback condition it almost never manages to establish a lasting interaction with agent B, its 'playback' partner (Fig. 5, right-hand side).

Given that the interaction process between the agents is normally highly robust even under very noisy environmental conditions, the detrimental effects of the removal of social contingency may be different from simply adding noise to the system. This failure indicates that the responsiveness of the other agent during the normal condition plays a role in the success of the interaction process. The playback condition with 0 noise may appear to be an exception to this rule, but this is not the case. In that kind of absolutely identical replay the behavior of the playback agent cannot be meaningfully distinguished from the interaction-dependent behavior of the original live agent.

To get an intuitive insight into the differences between the live and playback conditions, we can again visualize the time series of the trials. Figure 6 shows two representative playback trials, using the normal trials shown in Fig. 4 as the recording. For a more detailed behavioral analysis, please refer to Appendix 2.

The results of the behavioral tests show that the behavior of a simulated agent is sensitive to the other's responsiveness, just as was found in Murray and Trevarthen's original psychological study. In particular, responsiveness is needed in order for the simulated agents to coordinate a common direction of movement. However, sensitivity to the other's responsiveness alone is not sufficient to make a strong case for the extended body. In cognitivist psychology, for example, the behavioral manifestation

[10] Technical aside: the playback trials do differ from the live trials in one essential way, namely by use of a different random number sequence to generate the noise. If we used the same random number sequence, then there would be no discernable behavioral difference between a live trial and a playback trial since we are dealing with a deterministic system. See also Fig. 5 when the noise level is set to 0.

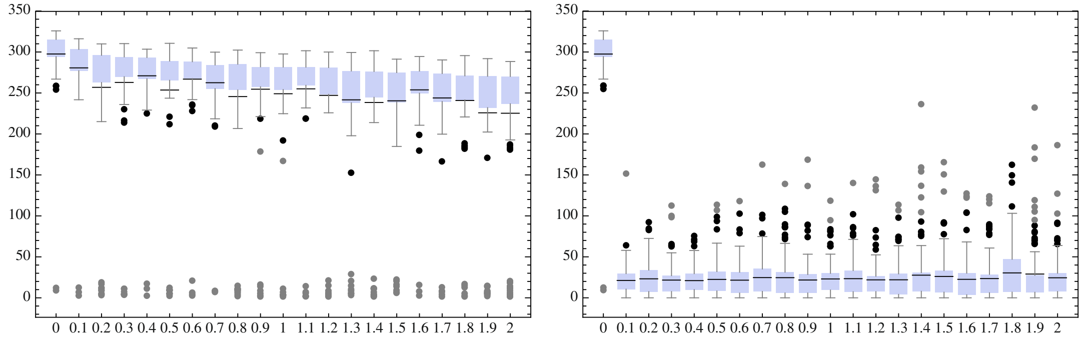


Fig. 5 Box-and-whisker summary of the scores of 21 sets of 100 trials for a range of noise variances (*x*-axis) shown for the 'live' condition (left figure) and the 'playback' condition (right figure). The scores on the *y*-axis are a measure of the distance traveled to the final point of contact. The original noise variance used during the evolutionary optimization process was 0.05. *Left:* Note the smooth decrease of average performance in the live condition, which continues to be highly successful even when the noise is increased 40 fold and more. *Right:* In contrast, in the playback condition the active agent fails to maintain its behavior at even the lowest noise levels. Since we are dealing with a deterministic system, setting the noise to 0 means that the dynamics of the playback trial will be absolutely identical to the live trial (and hence will receive the same score in both conditions)

of such 'social contingency' is simply taken as evidence for the existence of specialized cognitive 'modules' inside the heads of the individuals, which have the task of computing the codependent relationships between the individual's own and the other's behavior. In order to exclude this possibility in the case of the current model, and to determine whether an alternative explanation in terms of the extended body is viable, we have to analyze in more detail how the simulated agents' behaviors arise from the activity described by their dynamical systems. Demonstrating that such an alternative explanation is indeed valid for the case of this model has two consequences

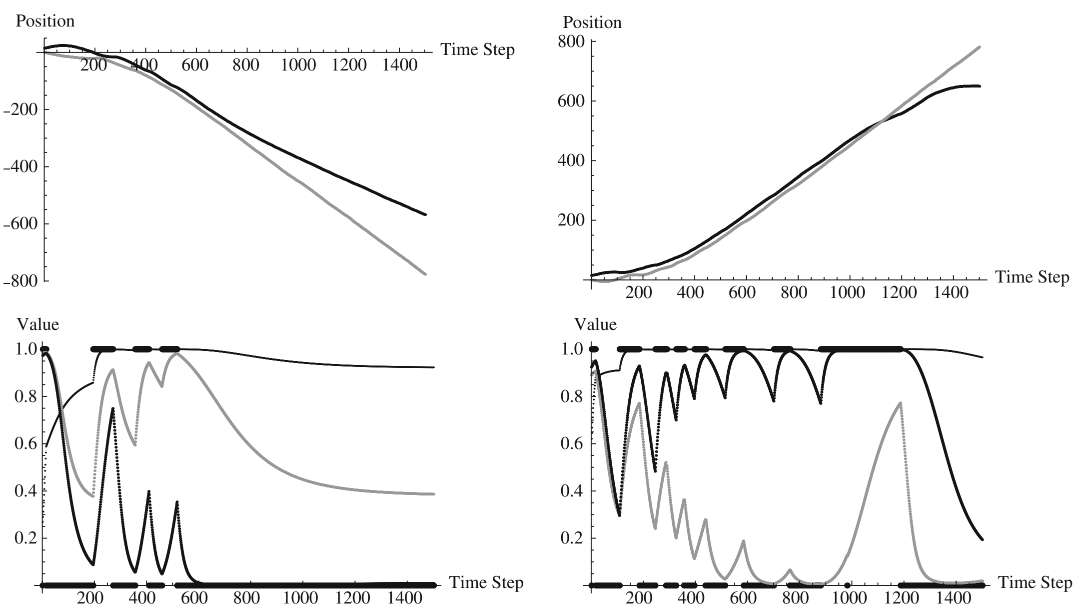


Fig. 6 Time series of the first 1500 time steps of two trials under 'playback' condition starting with the initial states of the two normal trials shown in Fig. 4. *First row:* Position of agent A (*black line*) and playback of agent B (*gray line*). *Second row:* Agent A's values of contact sensor (*thick black line*, at either 0 or 1), neuron 1 output (*thin black line*), neuron 2 output (*gray line*), and neuron 3 output (*black line*). Although not shown here, eventually the neurons of agent A settle into the same fixed-point equilibrium

for actual psychological experiments. It shows that sensitivity to social contingency on its own is not sufficient evidence for the traditional hypothesis of specialized ‘social contingency modules’ in the brain, and it opens up the possibility for interactive accounts based, for instance based on the notion of the extended body.

Dynamical analysis of the model agents

One helpful way of understanding the structure of the agent’s activity is to visualize how the states of its components would change (a vector) over a range of possible states (the state space). The result of this mathematical procedure is a vector field of the flows of activation, as illustrated in Fig. 7.

In contrast to the cognitivist assumption that the input layer is separated from the internal processing layer, it turns out that the two cannot be separated in this model. Notice how the organization of the vector field changes depending on the state of the input parameter (compare the left- and right-hand side of Fig. 7). The state of the agent’s bodily contact sensor has the capacity of modulating the way in which the internal activity of the agent is structured, and hence how its behavior is governed. Conversely, it is the interaction process that arises out of the behavior of the agents, and which determines how the status of the contact sensor changes over time. And this change in contact status in turn has the effect of guiding the internal dynamics of the agents into a transient region of state space that is inaccessible without mutually responsive bootstrapping (see Appendix 3).

The essential point of the analysis is that contact with the other agent does not merely serve as an external input to a pre-existing subpersonal architecture, as classical cognitive science would assume. It is the mutual parametric modulation between the agents, which has the effect of shaping the structure of the agent’s state space into a transient form that enables a task-relevant sensorimotor loop to emerge

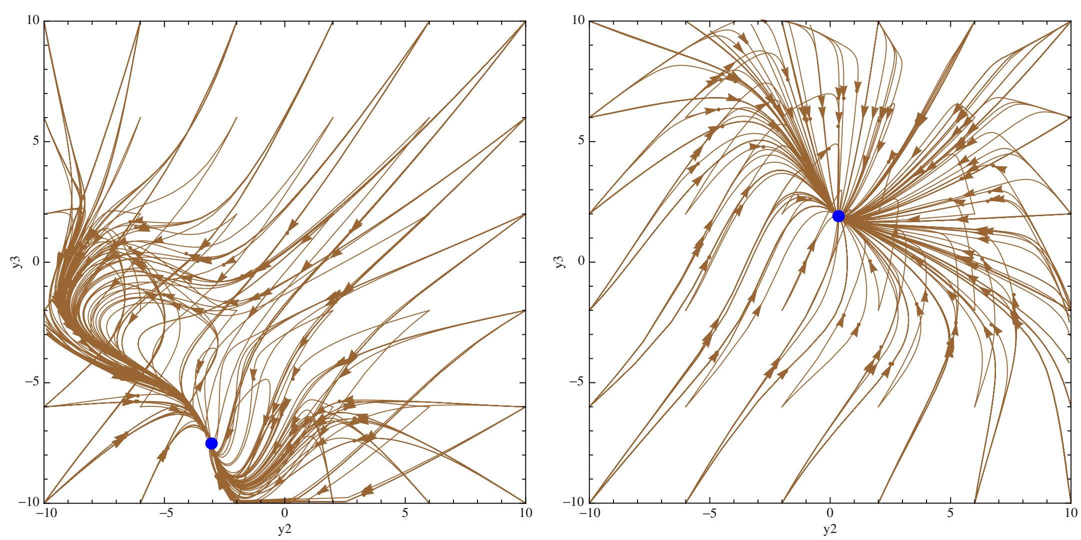


Fig. 7 Representative flow structure of a section of the activation space of the continuous-time recurrent neural network, when the input parameter is fixed to $I=0$ (left figure) and $I=1$ (right figure). Here we focus on the activation flows of neurons y_2 and y_3, omitting neuron y_1 for clarity. In both input conditions the flows of activation eventually converge on a single stable fixed-point attractor (*solid dot*). But the position of this equilibrium point and the structure of its basin of attraction depend on the state of the input parameter. The equilibria are located at (−3.05,−7.52) and (0.34, 1.90) for $I=0$ and $I=1$, respectively

and to be maintained. Here we thus have a basic dynamical account of the extended body: the behavior of the agents gives rise to an interaction process, which in turn transforms the agents' internal conditions of behavior generation, such that their behavior in fact becomes more suited to further sustain the interaction process, and so forth. Not only do the agents make use of each other's responsive presence to bootstrap into a more flexible behavioral domain, this process itself is a self-organized outcome of the self-maintaining interaction process that has emerged from their mutually responsive behavior. The bodies of the two simulated agents have become intertwined into one extended body through the interaction process.

Thus, in contrast to the claims of classical cognitive science, the model has demonstrated that behavioral sensitivity to social contingency in actual experimental situations does not in principle require complex cognitive modules based on inner representations or mental simulation routines. Instead, it is also possible that such social behavior is an interactive achievement based on the dynamical principles of bodily extension that we have analyzed. Moreover, given that the extended body is an emergent outcome that is distributed across two or more agents, it places less demands on the cognitive capacities of the individuals than the cognitivist account. It should therefore be considered as the default mechanism underlying social contingency in actual experimental situations, unless there is additional evidence, which shows that there are internal cognitive modules dedicated to this task.

Implications for body memory

So far we have mainly considered the role of inter-bodily resonance in the emergence of sensitivity to social contingency. The notion of an extended body memory was not made explicit in the dynamical analysis of the model agents' behavior, but it is relevant in at least two respects, namely when considering the initial origination and the subsequent maintenance of appropriate interactive motor skills.

First, it is noteworthy that the organization of the agents' behavior depends on which role each agent adopts. The overall direction of the mutual interaction must first be negotiated, and a different dynamic pattern or 'motor schema' is eventually acquired depending on this direction (see top row of Fig. 8 in Appendix 3). The origin of this organization cannot be ascribed to the actions of one individual alone; rather, it self-organizes out of the interaction process. The history of this process is embodied in the interactively shared organization of the two agents, and this 'inter-bodily memory' continues to be expressed behaviorally in terms of the role that an agent has converged on for a particular trial.

Second, what the dynamical analysis of the model also shows is that the persistence of acquired behavioral routines (as a transient pattern of activity) continues to depend on the actualization of further appropriate interactions (the interaction process); otherwise the 'implicit knowledge' decays away.[11] If we were to remove an agent's

[11] The phenomenological terms are used here in apostrophes to denote the corresponding processes on the systems level. The subjective experience and bodily dispositions involved in social interactions cannot be reduced to a systemic account of (robotic or other) agents. The dynamical systems view is able to model and analyze the structural or dynamical aspect of the interaction process, but not directly its lived, felt and co-experienced quality.

partner during the middle of a trial, the remaining agent would quickly become trapped in monotonous movement, as its transient dynamics can no longer be maintained and become absorbed by one of the fixed attractors (similar to what is shown in the bottom row of Fig. 8 in Appendix 3). Accordingly, in this case the 'body memory' of the agents was not only interactively acquired, it was also interactively maintained.

These dynamical insights into the model agents might shed a light on the problem of how implicit knowledge gets obtained and retained in natural agents. For instance, since the scientific tradition is dominated by a computationalist view of the mind, the issue of memory permanence appears not to be a problem of cognitive interest: the process of learning simply writes bits of knowledge into memory banks where they are stored and can be recalled at will. On this view, if there is a problem of memory storage at all, this is only a problem of the physical substrate. This essentially *static* view of memory storage hides the possibility that the persistence of implicit knowledge may actually be the result of a *dynamic* process. Accordingly, if we replace this static view of memory with a process-based conception, then cognitive science encounters the problematic of having to *explain* rather than being able to *assume* the persistence of implicit knowledge.

We note that this issue is not limited to empirical accounts of implicit memory. A similar worry can be raised about Husserl's [1893-1917] (1969) influential account of inner time-consciousness. It is worth quoting Husserl at length to show that neurophenomenology is not only about using phenomenology to improve empirical science, but also about using science to improve phenomenology, i.e. to establish a mutually informing dialogue.

> Retention is not a modification in which impressional data are really preserved, only in modified form: on the contrary, it is an intentionality—indeed, an intentionality with a specific character of its own. When a primal datum, a new phase, emerges, the preceeding phase does not vanish but is "kept in grip" (that is to say, precisely "retained"); and thanks to this retention, a looking-back at what has elapsed is possible. […] Each phase, by being retentionally conscious of the preceding phase, includes in itself the entire series of elapsed retentions in the form of a chain of mediate intentions: it is precisely in this way that duration-unities […] become constituted. (Husserl [1893-1917] (1969), Appendix IX)

At least on this particular phenomenological account of time consciousness there are no constraints to the size of the nested chain of distinct retentions, and there is no intrinsic mechanism of forgetting. Accordingly, each individual past moment always remains at least potentially accessible throughout the subject's lifetime. But this does not appear to be an accurate description of the memory of an embodied subject. Part of the problem is Husserl's reliance on the language of intentionality to describe the retention itself, which seems to entail that the past is retained as a static chain of discrete moments. If instead we use the terminology of a continuous dynamical system to describe time consciousness (e.g. Varela 1999), then the past can be conceived more like a holistic reservoir, and the continuous modification and degradation of memory thereby becomes an intrinsic aspect of temporality. This may be closer to

Husserl's later position (Summa 2011), and it opens up the important question of how memory structures can be retained in their specificity at all.[12]

In the case of the model agents the persistence of their 'implicit memory' is achieved by means of a practice-based 'use it or lose it' strategy. The agent's 'implicit memory' has an ongoing dependency on the existence of an appropriate context in which those memorized skills can be repeatedly realized and practiced. Of course, while this kind of volatile memory was sufficient for the requirements of the task, in a complex and changing environment it would be advantageous to effect a reorganization, which reduces this environmental dependency. What would be needed is a developmental ratchet effect, whereby skills acquired during interaction with others can be maintained on an individual basis, and perhaps could thereby even be deployed outside of the interaction context in which the original memory was embodied.[13] We speculate that this reorganization may be one of the goals of maturation in human development, particularly during adolescence.

Discussion

We have considered two aspects of the extended body, namely inter-bodily resonance and inter-bodily memory, from the perspectives of phenomenology, enactive cognitive science, and agent-based modeling. This neurophenomenological approach has enabled us to address the results of Murray and Trevarthen's psychological study in a novel way. We suggest that the negation of inter-bodily resonance with the mother during the playback condition is, on the one hand, experienced by the child as an irritating 'misfit' of the mutual contact, and, on the other hand, is expressed as a growing instability and eventual breakdown of the interaction process. By this twofold approach we arrive at a complementary account of one and the same interaction process as seen from a 1st/2nd person perspective and from a 3rd person perspective.

It is also noteworthy that the dynamical systems account, which mediates between these different perspectives, is neutral regarding whether the temporal changes that it describes pertain to an extended *lived* body or an extended *living* body. That is, it is neutral about whether it describes changes in the body that are distinguished from the first-person or third-person perspective. Interestingly, there is some structural isomorphism between the interactions taking place on the level of lived bodies and the interactions on the level of living bodies. In both cases, for instance, one can speak in terms of resonance and of its disturbance, of coordination and of its breakdown, as well as of mutual incorporation and of relative isolation. This indicates that there is an

[12] In addition, once it is accepted that the preservation of past retentions in their individual specificity is a dynamic achievement rather than a static given, we must approach the phenomenological problem of the constitution of self-consciousness from a different angle. This is because it is precisely due to "retention that consciousness can be made into an object" (Husserl, [1893-1917] (1969), Appendix IX).

[13] There is an important distinction between participatory sense-making among a group of undifferentiated agents and the act of making sense of others *as* others (Gallagher 2009). If we accept that the latter ability develops on the basis of the former, then something like this reorganization is also needed to explain our ability of direct perception of others even in the absence of any immediate interaction (i.e. the classical case of social cognition). See also the discussion by Stout (this issue).

underlying existential unity of a living-lived body before we dissect the phenomenon of life as either living or lived by studying it from the vantage points of different perspectives.

Analyzing embodied intersubjectivity in terms of an extended body in this manner has the advantage that, in contrast to theories based on theory and/or simulation, it does not inadvertently introduce an egocentric homunculus into the domains of pre-reflective experience and unconscious bodily activity. And yet at the same time it still manages to take the personal and subpersonal levels of description into account. Furthermore, the analysis of the model has provided us with a proof of concept that the notion of the extended body is more than merely a poetic or fanciful description of phenomenology. It can be cashed out in precise mathematical terms, which have the objectivity and mathematical rigor we have come to expect of the natural sciences. The concept of the extended body thus offers itself as the basis for an alternative scientific framework of social cognition, which can be explored by future empirical work in interaction studies. At the same time the concept's concrete foundation in phenomenology prevents the empirical findings and dynamical principles to be used in the service of functionalist reduction or physicalist elimination.

Conclusion

The scientific study of social cognition has shifted toward a greater consideration of the role of embodied intersubjectivity, which aligns it with a growing body of empirical and phenomenological findings. However, we have identified significant shortcomings in the mainstream accounts of embodied intersubjectivity that stem from a continuing reliance on the cognitivist framework of classical cognitive science. One major motivation for this persistent commitment to the core principles of cognitivism is derived from a confused phenomenology, whereby the states of other peoples' minds are assumed to be hidden from one's own perceptual experience, and consequently have to be secondarily obtained by means of some 'Theory of Mind' mechanism. It is therefore important to clarify the phenomenology of embodied intersubjectivity, such that the possibility of having direct perception of the presence of other minds is recognized as an experiential fact (see, e.g., Krueger 2012; Stout 2012).

However, in order to make further progress in this direction we must also recognize that classical cognitive science is mainly interested in explaining the functioning of the mind in terms of subpersonal representational cognitive architectures. The important task of phenomenological clarification is therefore in danger of simply being marginalized (e.g. Spaulding 2010), especially if it is not explicitly complemented by the development of a convincing alternative account of the neural and bodily events that are taking place at the subpersonal level. Accordingly, one aim of our neurophenomenological case study of the extended body was to provide the beginnings of such an alternative account in terms of dynamical systems theory. This mathematical framework allows us to integrate body, mind, and sociality in a phenomenologically-informed and non-reductive manner, and without the need for a subpersonal homunculus or representational architectures.

Finally, given the continuing prominence in cognitive science of the debate regarding the extended mind (Clark 2008), it is useful to give a brief indication of how our proposal of the extended body relates to this trend. In brief, it should be evident that if we accept that the mind is embodied, and that the dynamical entanglement of embodied intersubjectivity can give rise to an extended body, then the mind can also be extended in this socially embodied manner. However, whereas the extended mind literature is mainly interested in extending the bounds of cognition, we are also concerned with extending the bounds of experiencing. Clark (2009) is unconvinced that the conscious mind can be extended. But according to the neurophenomenological framework we have developed in this paper, we propose that the dynamical extension of 3rd-person processes of the living body during social interaction cannot be separated from a complementary experiential extension of the 1st-person perspective of the lived body. Going beyond the idea of an isolated extended conscious mind, we propose that the phenomenology of the 2nd-person perspective could be grounded in two dynamically intertwined living-lived bodies (Froese 2011). We suggest that it is on the basis of this kind of mutual entanglement that we can share and participate in each other's experience.

Acknowledgements The collaboration between Froese and Fuchs was initially enabled by the COST Action on Consciousness (BM0605), which funded a Short-Term Scientific Mission for Froese to work with Fuchs at the University of Heidelberg. During the later completion of this paper Froese was financially supported by a Grant-in-Aid awarded by the Japanese Society for the Promotion of Science (JSPS). We thank Rasmus Thybo Jensen and the anonymous reviewers for their detailed comments, and Takashi Ikegami for many helpful discussions. The software implementation of the agent-based model made use of Randall D. Beer's "Evolutionary Agents v1.1.2" C++ package. The mathematical analysis of the agent's dynamical system was performed with the help of Randall D. Beer's "Dynamica v1.0.4" Mathematica package.

Appendix 1: Additional modeling details

The standard format of the continuous-time recurrent neural network (CTRNN) equations is as follows:

$$\tau_i \dot{y}_i = -y_i + \sum_{j=1}^{N} w_{ji}\sigma\left(y_j + \theta_j\right) + m_i I \quad i = 1, \ldots, N$$
$$\sigma(x) = \frac{1}{(1+e^{-x})}$$

Following Beer (1995, 2003), the form of these equations can be interpreted in a rough analogy to aspects of a functioning neuron, where y_i is the state of the average membrane potential of the ith neuron, y_i denotes the rate of change of this state, τ_i is the neuron's membrane time constant, w_{ji} is the strength of the connection from the jth to the ith neuron, and θ_i is a bias term. The short-term average firing rate of the ith neuron is given by the output of the standard logistic activation function σ_i. Note, however, that this is only one possible interpretation of these equations, and it may be better to just think of them as describing a network of arbitrary dynamical components. The state of the input parameter I is equal to the state of the contact sensor (i.e. 0 for a lack of bodily contact, and 1 for the presence of bodily contact), where m_i is the magnitude of that state. Since the environment holds nothing but the two agents, the input I will always be equal for the two agents (either there is mutual overlapping or there is not).

On the basis of the output of the two functions σ_2 and σ_3 (range [0, 1] mapped into range [−1, 1]) we determine an agent's movement in the one-dimensional space. Rightward and leftward movements are modeled as an increase and decrease of position in 1D space, respectively. The total change of position (velocity) is taken to be the difference between the respective contributions of the two dynamical components, and the contribution of each component v_i is calculated by adding to its output σ_i a stochastic element $Noise_i$, i.e. small random number drawn from a Gaussian distribution [mean=0, variance=0.05], and then scaling the result by an output gain g_i.

$$v_i = g_i(\sigma_i + Noise_i)$$

$$velocities = \begin{cases} v_3 - v_2, & Agent\ A \\ v_2 - v_3, & Agent\ B \end{cases}$$

The noise term introduces a small element of uncertainty into the model, such that the agents cannot rely on a direct mapping between their 'expected' change in position and their actual change in position. In contrast to the original work by Froese and Di Paolo (2008) we did not introduce sensor noise. Note that the total velocity of each agent is calculated slightly differently, with the two components v_2 and v_3 having the opposite sign. This represents the fact that the agents face each other in the one-dimensional environment, and their directions of movement are accordingly reversed.

Details of the evolutionary algorithm that was used in order to optimize the configuration of the CTRNNs can be found in Froese and Di Paolo (2008). Essentially, the probability of a particular configuration being chosen for the next round of optimization was set to be proportional to the agents' ability to maximize the distance traveled together. This ability was objectively measured in terms of the distance traveled from their point of origin (the agents start at a random location within range [−25, 25] of each other) to their last point of contact before the end of the trial (after 50 units of time). The optimization process was terminated as soon as a configuration was found that enabled the agents to consistently coordinate their movements over a variety of initial conditions. Note that for the current study we did not repeat this optimization procedure, but simply took the particular configuration that was found by Froese and Di Paolo as our starting point.[14] While this configuration is not the only way to solve the coordination task, it serves as a fitting proof of concept.

Appendix 2: Additional behavioral analysis

This Appendix provides a more detailed assessment of the behavioral interaction of the agents. The time series tracing the change of position of the agents over time

[14] The CTRNN's parameter values (rounded to 3 decimal places) are the following: biases θ_1=−0.827, θ_2=2.567, θ_3=2.930; time constants τ_1=1.209, τ_2=1.581, τ_3=1.137; weights $w_{1,1}$=3.303, $w_{2,1}$=0.644, $w_{3,1}$=0.205, $w_{1,2}$=−3.660, $w_{2,2}$=1.0479, $w_{3,2}$=−7.920, $w_{1,3}$=−5.803, $w_{2,3}$=−5.768, $w_{3,3}$=2.334; input gains m_1=m_2=m_3=10.861; output gains g_1=24.875, g_2=44.512.

(Fig. 4, top row) enables us to observe at least three characteristic features of the interaction.

First, the interaction process can be divided into two phases. During the first 500 hundred time steps the agents move apart and closer again with a decreasing magnitude of oscillation. The future direction is already discernable as a bias in the initial oscillations (but it can always still be affected by noise). When the agents settle on a preferred distance from each other, which happens to be just at the margin of making contact, the oscillations continue with reduced magnitude and the shared velocity is increased. At this point the second phase starts; the direction of the interaction pattern for the remainder of the trial has been established.

Second, the agents appear to exhibit an example of ‘active perception’. The informational value of the contact sensor in itself is very limited; it only provides a single binary signal (on/off). Accordingly, prolonged contact is not beneficial because it makes it uncertain to what extent the agents are actually overlapping, and prolonged absence of contact makes it uncertain how far they have drifted apart. By maintaining only a transient, oscillatory contact at the margin of the sensor boundary, the agents therefore ensure that the sensor signal provides accurate information about the relative position of the agents.

Third, it turns out that the relative position between the agents is always the same in the second phase of a trial, no matter whether they eventually end up moving leftwards or rightwards. This has the effect of decreasing the complexity of the task by turning four distinct possibilities (i.e. two types of relative position and two types of direction) into two possibilities (two types of direction). It also further enhances the informational value of the contact sensor, because while coupling could take place at one of two sides of the body (depending on relative position), it now is practically arranged to always take place at one and the same side only. Nevertheless, the task still remains nontrivial: a loss of contact has two different implications depending on context: in one case reestablishing contact requires a decrease of velocity in the shared direction (a ‘leader’ has to fall back), while in the other it requires an increase of velocity in the shared direction (a ‘follower’ has to catch up). In other words, the coordination of a common direction of movement is also about converging on complementary roles, of who will lead and who will follow.

Further behavioral investigations have revealed the following results. If the design of the playback condition is modeled more closely on the one originally used by Murray and Trevarthen, namely that recording of behavior only starts after the interaction has already been established and the playback is then started in the middle of the trial, we find that the interaction process still ends up breaking down. However, it only breaks down if the noise level was set to a value higher than 0, otherwise the live agent manages to maintain its behavior in the absence of the other’s responsiveness. At first sight this seems to be the same result as when we start playback from the beginning of the whole trial with 0 noise (see Fig. 5), but this is not the case. The data shown for the whole trial playback condition was derived by setting the internal activation state of the live agent to be the same it had during the normal condition. The middle of trial playback condition, on the other hand, does not reset the agent’s activation state to the same state it was in when the recording started. Interestingly, when the whole trial playback condition is started with 0 noise, but this time with newly randomized initial conditions for the live agent, then no interaction process is established.

The upshot of these results is that the role of the contingent responsiveness of the other agent changes during the progression of the trial. During the beginning of the trial the responsiveness is needed in order to coordinate the internal activation state of the agents. This makes sense because the agents need to converge on one or the other direction. After this initial coordination has taken place, the responsiveness of the other is needed to ensure that any discrepancies in coordinated movement that arise due to external perturbations do not accumulate over time to an unmanageable extent. Both of these roles of the other's responsiveness modify the conditions in which the behaviors of the agents are generated. But while joint noise regulation pertains to conditions of the environment, the coordination of internal activation states during the initial moments of a trial pertains to conditions involving the agents themselves. We are particularly interested in the latter case because the notion of the extended body posits that inter-bodily resonance can transform the internal milieu of the agents. Our dynamical analysis is therefore focused on events occurring in the initial stages of the trials only.

Appendix 3: Additional dynamical analysis

For a strong notion of the extended body, it is not sufficient that the agents have the capacity for causing a switch in each other's internal organization. This switch can be achieved by any kind of contact, in principle even by contact with non-agent objects that happen to cross the sensor. The idea of the extended body, on the other hand, requires that internal re-organization and mutual responsiveness are co-dependent factors, and that both of them are needed to establish coordinated movement. To put it differently, we are interested in (inter-bodily) socially contingent (intra-bodily) re-organization. In order to determine whether this is indeed the case in our model we must take a closer look at how an agent's activation changes during its ongoing interaction with the other agent. Figure 8 illustrates how the inter-bodily interaction process relates to the intra-bodily switching between the two possible flow structures and their divergent equilibrium points.

The zigzag pattern noticeable in all of the graphs in Fig. 8 stems from the repeated on-off switching of the contact sensor during the interaction between the agents. Depending on the sensor status, only one or the other of the equilibrium points is actually present and is attracting the state's trajectory. We can observe a ratchet effect that prevents the agents' state from merely oscillating between the same two locations. This is due to the specific flow structure of the equilibrium points' basins of attraction, which attracts the state in a nonlinear fashion (see Fig. 7). Note that when the interaction breaks down during the playback condition, the contact sensor remains off, the zigzag pattern disappears, and the agents' state finally settles into the attractor that is defined by input $I=0$.

As we know from Fig. 7 already, when the agents are on their own, i.e. without any change in input, the agents are fixed by the flow structure determined by $I=0$, which

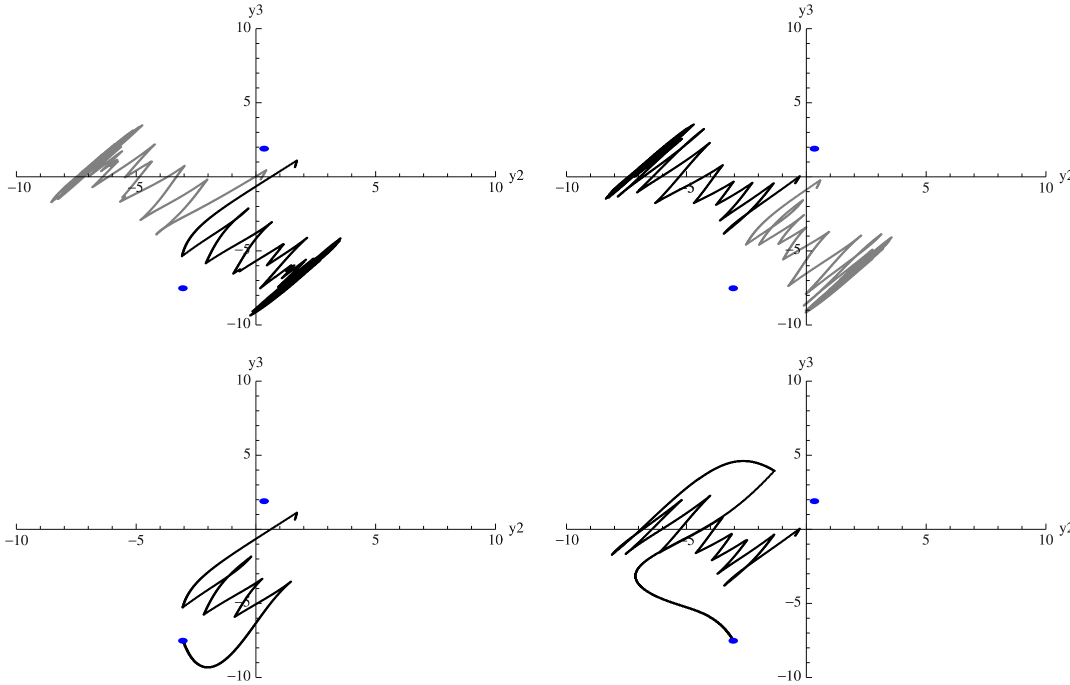


Fig. 8 The two agents' initial trajectories in state space for representative trials of the 'live' condition (*top row*, same trials as shown in Fig. 4), and those of agent A in the 'playback' condition (*bottom row*, same trials as shown in Fig. 6). Activations of agent A (*black line*) and agent B (*gray line*) always start near the origin and then progress outwards with time. The locations of the normally separately existing equilibrium points shown in Fig. 7 are superimposed simultaneously (*solid dots*). Note that during the live interaction the agents' activations are maintained as far-from-equilibrium transients, and that these can take place in two distinct regions of state space depending on whether the agents are moving leftward (*left column*) or rightward (*right column*). In contrast, agent A's activations always quickly converge to one of the equilibrium points during the playback conditions

is defined by a single equilibrium point, and which limits their behavior to movement in the same direction at a constant speed. What Fig. 8 reveals is that the robustness of the agents' behavior, as demonstrated by their resilience to external perturbations (Fig. 5), depends on their internal state following a transient pattern within a far-from-equilibrium region of state space. In this transient region the agent's internal flow structure effectively operates as a stable quasi-periodic equilibrium, rather than as a fixed-point attractor, which in this case has the desirable effect of expanding the agent's behavioral repertoire. Instead of being effectively limited to a single direction of movement, the agent can now move both left and right in a flexible manner due to the interactively stabilized transient region.

Importantly, while this transient region enables the kind of flexible behavior that is required for sustaining a responsive interaction, it is also the case that what enables an agent's internal state to first enter into (and then to remain within) this region is precisely the responsive behavior between the agents. The agents must guide each other's internal state into this transient region by switching each other's internal flow structure in an appropriate manner. This finding confirms that we are indeed dealing with a model of an extended body: *each of the agent's intra-bodily dynamics is extended by the other agent's intra-bodily dynamics by means of the inter-bodily dynamics of their interaction process*.

References

Agnew, Z. K., Bhakoo, K. K., & Puri, B. K. (2007). The human mirror system: a motor resonance theory of mind-reading. *Brain Research Reviews, 54*, 286–293.

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

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*(1), 37–46.

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

Beer, R. D. (2000). Dynamical approaches to cognitive science. *Trends in Cognitive Sciences, 4*(3), 91–99.

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

Boden, M. A. (2006). Of islands and interactions. *Journal of Consciousness Studies, 13*(5), 53–63.

Carruthers, P., & Smith, P. K. (Eds.) (1996). *Theories of theories of mind*. Cambridge University Press.

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

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

Clark, A. (2009). Spreading the joy? Why the machinery of consciousness is (probably) still in the head. *Mind, 118*(472), 963–993.

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., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences, 14*(10), 441–447.

De Vignemont, F. (2004). The co-consciousness hypothesis. *Phenomenology and the Cognitive Sciences, 3*, 97–114.

Di Paolo, E. A. (2009). Editorial: the social and enactive mind. *Phenomenology and the Cognitive Sciences, 8*(4), 409–415.

Di Paolo, E. A., Noble, J., & Bullock, S. (2000). Simulation models as opaque thought experiments. In M. A. Bedau, J. S. McCaskill, N. H. Packard, & S. Rasmussen (Eds.), *Artificial life VII: Proc. of the 7th Int. Conf. on artificial life* (pp. 497–506). Cambridge: The 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.

Di Paolo, E. A., Rohde, M., & De Jaegher, H. (2010). 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: The MIT Press.

Downing, G. (2004). Emotion, body and parent–infant interaction. In J. Nadel & D. Muir (Eds.), *Emotional development: Recent research advances* (pp. 429–449). Oxford: Oxford University Press.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.

Fodor, J. A. (1992). A theory of the child's theory of mind. *Cognition, 44*, 283–296.

Froese, T. (2011). Breathing new life into cognitive science. *Avant. The Journal of the Philosophical-Interdisciplinary Vanguard, 2*(1), 113–129.

Froese, T., & Di Paolo, E. A. (2008). Stability of coordination requires mutuality of interaction in a model of embodied agents. In M. Asada, J. C. T. Hallam, J.-A. Meyer, & J. Tani (Eds.), *From animals to animats 10: Proc. of the 10th Int. Conf. on simulation of adaptive behavior* (pp. 52–61). Berlin: Springer.

Froese, T., & Di Paolo, E. A. (2010). Modeling social interaction as perceptual crossing: an investigation into the dynamics of the interaction process. *Connection Science, 22*(1), 43–68.

Froese, T., & Gallagher, S. (2010). Phenomenology and artificial life: toward a technological supplementation of phenomenological methodology. *Husserl Studies, 26*(2), 83–106.

Froese, T., & Di Paolo, E. A. (2011a). The enactive approach: theoretical sketches from cell to society. *Pragmatics & Cognition, 19*(1), 1–36.

Froese, T., & Di Paolo, E. A. (2011b). *Toward minimally social behavior: Social psychology meets evolutionary robotics.* In G. Kampis, I. Karsai, & E. Szathmáry (Eds.), *Advances in artificial life. Darwin meets von Neumann. 10th European Conference, ECAL 2009* (pp. 426–433). Berlin: Springer-Verlag.

Froese, T., Gould, C., & Barrett, A. (2011). Re-viewing from within: a commentary on first- and second-person methods in the science of consciousness. *Constructivist Foundations, 6*(2), 254–269.

Fuchs, T. (1996). Leibliche Kommunikation und ihre Störungen. *Zeitschrift fuer Klinische Psychologie Psychopathologie und Psychotherapie, 44*, 415–428.

Fuchs, T. (2008). *Leib und Lebenswelt: Neue philosophisch-psychiatrische essays*. Kusterdingen: Die Graue Edition.

Fuchs, T. (2012). The phenomenology of body memory. In S. C. Koch, T. Fuchs, M. Summa, & C. Müller (Eds.), *Body memory, metaphor and movement* (pp. 9–22). Amsterdam: John Benjamins.

Fuchs, T., & De Jaegher, H. (2009). Enactive intersubjectivity: participatory sense-making and mutual incorporation. *Phenomenology and the Cognitive Sciences, 8*(4), 465–486.

Gallagher, S. (1997). Mutual enlightenment: recent phenomenology in cognitive science. *Journal of Consciousness Studies, 4*(3), 195–214.

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

Gallagher, S. (2007). Simulation trouble. *Social Neuroscience, 2*(3–4), 353–365.

Gallagher, S. (2008). Direct perception in the intersubjective context. *Consciousness and Cognition, 17*(2), 535–543.

Gallagher, S. (2009). Two problems of intersubjectivity. *Journal of Consciousness Studies, 16*(6–8), 289–308.

Gallese, V. (2005). Embodied simulation: from neurons to phenomenal experience. *Phenomenology and the Cognitive Sciences, 4*, 23–48.

Gallese, V. (2007). Before and below 'theory of mind': embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society B, 362*, 659–669.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences, 8*(9), 396–403.

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: Lawrence Erlbaum.

Giunti, M. (1995). Dynamical models of cognition. In R. F. Port & T. van Gelder (Eds.), *Mind in motion: Explorations in the dynamics of cognition* (pp. 549–572). Cambridge: The MIT Press.

Goldman, A., & de Vignemont, F. (2009). Is social cognition embodied? *Trends in Cognitive Sciences, 13* (4), 154–159.

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.

Heath, J. (2011). Methodological individualism. In: E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Spring 2011 Edition)*, URL = http://plato.stanford.edu/archives/spr2011/entries/methodological-individualism/.

Husserl, E. [1893-1917] (1969). Zur Phänomenologie des inneren Zeitbewusstseins (1893-1917). In: *Husserliana. Gesammelte Werke, Band 10*. Berlin: Springer

Hutchins, E. (1995). *Cognition in the wild*. Cambridge: The MIT Press.

Hutto, D. (2008). *Folk psychological narratives: The sociocultural basis of understanding reasons*. Cambridge: MIT Press.

Hutto, D., & Ratcliffe, M. (Eds.). (2010). *Folk psychology re-assessed*. Dordrecht: Springer.

Iizuka, H., & Di Paolo, E. A. (2007). Minimal agency detection of embodied agents. In F. Almeida e Costa, L. M. Rocha, E. Costa, I. Harvey, & A. Coutinho (Eds.), *Advances in artificial life: Proc. of the 9th Euro. Conf. on artificial life* (pp. 485–494). Berlin: Springer.

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

Jackson, P. L., & Decety, J. (2004). Motor cognition: a new paradigm to study self-other interactions. *Current Opinion in Neurobiology, 14*, 259–263.

Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends in Cognitive Sciences, 9*(1), 21–25.

Keysers, C., & Fadiga, L. (2008). The mirror neuron system: new frontiers. *Social Neuroscience, 3*(3–4), 193–198.

Krueger, J. (2012). Seeing mind in action. *Phenomenology and the Cognitive Sciences*.

Lutz, A. (2002). Toward a neurophenomenology of generative passages: a first empirical case study. *Phenomenology and the Cognitive Sciences, 1*, 133–167.

Lutz, A., & Thompson, E. (2003). Neurophenomenology: integrating subjective experience and brain dynamics in the neuroscience of consciousness. *Journal of Consciousness Studies, 10*, 31–52.

Lyons-Ruth, K., Bruschweiler-Stern, N., Harrison, A. M., Morgan, A. C., Nahum, J. P., Sander, L., Stern, D. N., & Tronick, E. Z. (1998). Implicit relational knowing: its role in development and psychoanalytic treatment. *Infant Mental Health Journal, 19*, 282–289.

Merleau-Ponty, M. [1945]. *Phénomènologie de la perception*, trans. by: C. Smith (1962), *Phenomenology of perception*. New York: Routledge & Kegan Paul.

Merleau-Ponty, M. (1960). “Les relations avec autrui chez l’enfant”, Paris, France: *Cours de Sorbonne*, trans. by W. Cobb, “The Child’s Relations with Others”. In: M. Merleau-Ponty (1964), *The Primacy of Perception And Other Essays on Phenomenological Psychology, the Philosophy of Art, History and Politics*, J. M. Edie (Ed.), (pp. 96-155). Evanston: Northwestern University Press.

Morganti, F., Carassa, A., & Riva, G. (Eds.). (2008). *Enacting intersubjectivity: A cognitive and social perspective on the study of interactions*. Amsterdam: Ios Press.

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: Alex.

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.

Noë, A. (2004). *Action in perception*. Cambridge: The MIT Press.

Petitmengin, C., Navarro, V., & Le Van Quyen, M. (2007). Anticipating seizure: pre-reflective experience at the center of neuron-phenomenology. *Consciousness and Cognition, 16*(3), 746–764.

Port, R. F., & van Gelder, T. (Eds.). (1995). *Mind as motion: Explorations in the dynamics of cognition*. Cambridge: MIT Press.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(4), 515–526.

Quinn, M., Smith, L., Mayley, G., & Husbands, P. (2003). Evolving controllers for a homogeneous system of physical robots: structured cooperation with minimal sensors. *Philosophical Transactions of the Royal Society of London A, 361*, 2321–2343.

Ratcliffe, M. (2007). *Rethinking commonsense psychology: A critique of folk psychology, theory of mind and simulation*. Hampshire: Palgrave Macmillan.

Reddy, V. (2008). *How infants know minds*. Cambridge: Harvard University Press.

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., Fogassi, L., & Gallese, V. (1999). Resonance behaviors and mirror neurons. *Archives Italiennes de Biologie, 137*, 85–100.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience, 2*, 661–670.

Roy, J.-M., Petitot, J., Pachoud, B., & Varela, F. J. (1999). Beyond the gap: An introduction to naturalizing phenomenology. In J. Petitot, F. J. Varela, B. Pachoud, & J.-M. Roy (Eds.), *Naturalizing phenomenology: Issues in contemporary phenomenology and cognitive science* (pp. 1–80). Stanford: Stanford University Press.

Spaulding, S. (2010). Embodied cognition and mindreading. *Mind & Language, 25*(1), 119–140.

Stawarska, B. (2006). Introduction: intersubjectivity and embodiment. *Phenomenology and the Cognitive Sciences, 5*(1), 1–3.

Starwaska, B. (2009). *Between you and I. Dialogical phenomenology*. Ohio: Ohio University Press.

Stern, D. N. [1985] (1998). *The interpersonal world of the infant: A view from psychoanalysis and developmental psychology*. New York: Basic Books.

Stewart, J., Gapenne, O., & Di Paolo, E. A. (Eds.). (2010). *Enaction: Towards a new paradigm for cognitive science*. Cambridge: MIT Press.

Stich, S., & Nichols, S. (1991). *Folk psychology: Simulation or tacit theory*. Rutgers University.

Stout, R. (2012). What someone’s behaviour must be like if we are to be aware of their emotions in it. *Phenomenology and the Cognitive Sciences*.

Stuart, S. (2011), Enkinaesthesia, Biosemiotics and the Ethiosphere. *Cybernetics & Human Knowing*, forthcoming.

Summa, M. (2011). Das Leibgedächtnis. Ein Beitrag aus der Phänomenologie Husserls. *Husserl Studies, 27*, 173–196.

Thelen, E., & Smith, L. B. (1994). *A dynamical systems approach to the development of cognition and action*. Cambridge: The MIT Press.

Thompson, E. (2001). Empathy and consciousness. *Journal of Consciousness Studies, 8*(5–7), 1–32.

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

Torrance, S. (2009). Contesting the concept of consciousness. *Journal of Consciousness Studies, 16*(5), 111–126.

Torrance, S., & Froese, T. (2011). An inter-enactive approach to agency: phenomenology, dynamics, and sociality. *Humana Mente, 15*, 21–53.

Trevarthen, C. B. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Ed.), *Before speech* (pp. 321–347). Cambridge: Cambridge University Press.

Trevarthen, C. (2005). First things first: infants make good use of the sympathetic rhythm of imitation, without reason or language. *Journal of Child Psychotherapy, 31*(1), 91–113.

van Gelder, T. (1999). Wooden iron? Husserlian phenomenology meets cognitive science. In J. Petitot, F. J. Varela, B. Pachoud, & J.-M. Roy (Eds.), *Naturalizing phenomenology: Issues in contemporary phenomenology and cognitive science* (pp. 245–265). Stanford: Stanford University Press.

Varela, F. J. (1996). Neurophenomenology: a methodological remedy for the hard problem. *Journal of Consciousness Studies, 3*(4), 330–349.

Varela, F. J. (1997). The naturalization of phenomenology as the transcendence of nature: searching for generative mutual constraints. *Alter: Revue de Phénoménologie, 5*, 355–381.

Varela, F. J. (1999). The specious present: a neurophenomenology of time consciousness. In J. Petitot, F. J. Varela, B. Pachoud, & J.-M. Roy (Eds.), *Naturalizing phenomenology: Issues in contemporary phenomenology and cognitive science* (pp. 266–317). Stanford: Stanford University Press.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge: The MIT Press.

Wheeler, M. (2005). *Reconstructing the cognitive world: The next step*. Cambridge: The MIT Press.

Zahavi, D. (2001). Beyond empathy: phenomenological approaches to intersubjectivity. *Journal of Consciousness Studies, 8*(5–7), 151–167.

Zlatev, J., Racine, T. P., Sinha, C., & Itkonen, E. (Eds.) (2008). *The shared mind: Perspectives on intersubjectivity*. John Benjamins Publishing Company.