Searching for the conditions of genuine intersubjectivity: From agent-based models to perceptual crossing experiments

Tom Froese [1, 2]

[1] Institute for Applied Mathematics and Systems Research (IIMAS), National Autonomous University of Mexico (UNAM), Mexico City, Mexico
[2] Center for the Sciences of Complexity (C3), UNAM, Mexico City, Mexico
Correspondence: t.froese@gmail.com

Abstract

Enactivists are searching for the conditions of genuine intersubjectivity. Theory of mind approaches to social cognition have come a long way from folk psychological theorizing by paying more attention to neuroscientific evidence and phenomenological insights. This has led to hybrid accounts that incorporate automatic processing and allow an instrumental role for perception and interaction. However, two foundational assumptions remain unquestioned. First, the *cognitive unconscious*: explanations assume there is a privileged domain of subpersonal mechanisms that operate in terms of representational personal-level concepts (belief, desire, inference, pretense, etc.), albeit unconsciously. Second, *methodological individualism*: explanations of social capacities are limited to mechanisms contained within the individual. The enactive approach breaks free from these representationalist-internalist constraints by integrating personal-level phenomenology with multi-scale dynamics occurring within and between subjects. This formal and empirical research on social interaction supports the possibility of *genuine intersubjectivity*: we can directly participate in the unfolding of each other's experience.

Keywords: theory of mind, social cognition, cognitive unconscious, experience, enactive approach, subpersonal mechanisms, phenomenology, social interaction, methodological individualism, intersubjectivity.

1. Introduction

1.1 Asocial assumptions at the origins of social cognition research

The existence of other subjects apart from oneself has long been a challenging topic in analytic and phenomenological philosophy, the so-called problem of other minds (for recent reviews of the literature, see Hyslop, 2016; Overgaard, 2012). The same is true of cognitive science: for if my first-person experience is only realized by processes within my individual body, or even just in my brain, my experience of the existence of other minds must also result from some of those internal processes. In particular, this leads to the mainstream hypothesis that my experience of another person must therefore be constituted by a mental representation created by processes in my brain, for example by inference and/or simulation. This standard working hypothesis is often referred to as *methodological individualism*, a phrase that has been adapted from sociology to signify the sufficiency of individual processes for explaining social cognition even in the context of interaction with others (Boden, 2006).

The details of this representationalist-internalist response to the problem of other minds, also known as *theory of mind*, vary substantially. But most accounts implicitly agree that whether other minds exist or not – and most if not all in the end assume that other minds do exist – is actually beside the point, at least for scientific explanatory purposes. Somewhat paradoxically, the aim is instead to explain social cognition in terms of one individual's isolated mind, such that the role played by other people is relegated to being an instrumental trigger for or external cause of social cognition mechanisms in that individual's brain. Hence, these theory of mind accounts also agree that social interaction between people cannot be *constitutive* of social cognition, i.e. the mechanisms of social cognition cannot be directly realized in terms of one's interaction with others.

To be fair, it is becoming widely accepted that social interaction may call forth special forms of shared or we-intentionality. But even so, such an "irreducibly collective mode" (p. 160) is generally taken to be realized within one individual's head (Gallotti & Frith, 2013). In other words, even when it is accepted that jointly acting according to a "we intend to do" is qualitatively distinct from and cannot be reduced to a mere sum of several "I intend to do", this concession of collective intentionality does not challenge methodological individualism, as was famously emphasized by Searle:

> Of course I take it in such cases that my collective intentionality, is in fact shared; I take it in such cases that I am not simply acting alone. But I could have all the intentionality I do have even if I am radically mistaken, even if the apparent presence and cooperation of other people is an illusion, even if I am suffering a total hallucination, even if I am a brain in a vat. Collective intentionality in my head can make a purported reference to other members of a collective independently of the question whether or not there actually are such members. (Searle, 1990, p. 407)

For Searle, and according to methodological individualism more generally, it is sufficient that my mind/brain behaves *as if* there are other agents with minds of their own in my environment, but whether they really exist makes no difference for explaining the phenomenon of collective intentionality. That is, even though traditional accounts, as epitomized by Searle's proposal, generally accept that other minds exist in addition to our own, their methodological individualism means that these accounts are similarly consistent with the extreme possibility of solipsism. They are essentially neutral on this important ontological question (whereas this kind of Cartesian skepticism is incompatible with the alternative claim that interaction with others can be constitutive of collective intentionality and of certain forms of social cognition). Of course, no one takes the possibility of solipsism seriously. Yet this problematic 'as if' assumption about the existence of other minds tends to motivate the type of explanations offered by theory of mind approaches: what ultimately matters is that my intentional stance is useful when I am dealing with certain complex 'objects' in my environment by allowing me to better explain and predict their future movements. Dennett has formulated a canonical version of this type of explanation:

> Here is how it works: first you decide to treat the object whose behavior is to be predicted as a rational agent; then you figure out what beliefs that agent ought to have, given its place in the world and its purpose. Then you figure out what desires it ought to have, on the same considerations, and finally you predict that this rational agent will act to further its goals in the light of its beliefs. A little practical reasoning from the chosen set of beliefs and desires will in most instances yield a decision about what the agent ought to do; that is what you predict the agent will do. (Dennett, 1987, p. 17)

Readers not trained in analytic philosophy or classical cognitive science may be forgiven for thinking how utterly strange and far removed from our normal lives these claims are. Searle's total hallucination and Dennett's autistic detachment certainly do not do justice to the complex phenomenology of normal human sociality in which we directly experience each other as persons in their own right (Ratcliffe, 2007). It is beyond the scope of this chapter to uncover the reasons for this rather strange state of affairs in cognitive science. Suffice it to note that these approaches follow rather straightforwardly from the assumption of mind-body dualism, which entails that we can perceive nothing but surface behavior whose causes are not directly given (Reddy, 2008). Moreover, we can understand the motivation for theory of mind accounts when we consider that traditionally the function of perception has been limited to making internal representations of physical objects located in the external environment. For if the mind is isolated inside the head, and if only the other's object-body is in principle accessible by means of my perception, then it must require extra cognitive work to ascertain others' mental states (i.e. a theory of mind).

This assumption, often known as the assumption of hidden minds, has been criticized on phenomenological grounds (e.g., Gallagher, 2012; but see Bohl & Gangopadhyay, 2014). The problem is that an encounter of others as opaque objects whose behavior must be inferred and predicted may describe what it is like to suffer from certain kinds of psychopathology (Froese, Stanghellini, & Bertelli, 2013), but it is not our normal condition with most of the people we meet everyday. It would therefore be preferable to develop an approach that is more in line with actual experience, that is, which agrees with the fact that we always already find ourselves in a world shared with others, while at the same time staying scientifically grounded in formal models and empirical results.

1.2 Phenomenological clarification of social cognition

As long noted by phenomenological philosophy, we have the direct perceptual experience that others exist, and we often share experiences with others while we engage in affiliative social interaction (León & Zahavi, 2016). Moreover, the latter types of intersubjective encounters are more adequately described as two first-person perspectives temporarily becoming integrated into a second-person perspective (Zahavi, 2016): you and me are having a laugh, we are eating a meal together, we are shaking hands, etc. We share an awareness of jointly being in one and the same unfolding situation. I will refer to these forms of being related by participating in each other's experience as *genuine intersubjectivity* in order to highlight that what I am interested in is that the other person as such plays a constitutive role in the social phenomenon to be explained This term also serves to thereby distinguish it from theory of mind approaches that are at least willing to accept that the phenomenology of these intimate forms of intersubjectivity is presenting us with shared moments of experience, but who nevertheless insist on explaining it within the constraints of methodological individualism.

Although in what follows I will not explicitly argue against the other foundational assumption of theory of mind approaches, i.e. the cognitive unconscious, I will demonstrate that it is actually possible to investigate the conditions of genuine intersubjectivity without postulating any subpersonal explanations that appeal to representational concepts. Indeed, I will not even make use of belief-desire psychology at the personal level because, arguably, for the kind of social interactions we will be considering, mutually engaging in embodied practices is sufficient for adequate social understanding (Kiverstein, 2011). We will consider the most minimal form of understanding that can be considered as social, namely the understanding that one is dealing with another subject.

This phenomenological tradition of emphasizing the varied possibilities of direct perceptual and embodied interactive understanding of other minds continues to be developed by contemporary phenomenological and enactive approaches to cognitive science (Gallagher, 2008; Froese & Leavens, 2014; Froese & Gallagher, 2012; Krueger, 2012). More recently, even among mainstream approaches there is growing acceptance that perceptual experience does not merely present us with meaningless surface behavior, but that we can directly perceive aspects of other minds (Wiltshire, Lobato, McConnell, & Fiore, 2015). The precise nature of the mechanisms that enable the direct perception of other minds, and the extent of this phenomenon, is currently hotly debated (Michael & De Bruin, 2015).

The problem is that it is one thing to agree that social encounters, like sharing a meal, are qualitatively experienced as social events, meaning that we can directly perceive the participation of another subject. But it is quite another to argue that the other person genuinely takes part in the constitution of that experience (De Jaegher, Di Paolo, & Adolphs, 2016). In other words, it is possible to accept the general validity of the phenomenological description, while insisting that it is still best explained by representationalist-internalist theory of mind mechanisms at the subpersonal level of one individual agent (Spaulding, 2010). For example, the discovery of mirror neurons lent itself to an account of direct perception in terms of simulation theory (Gallese, 2005). But even quite traditional 'theory' theory of mind approaches to direct social perception are possible (Carruthers, 2015), with some of them featuring new ways of thinking about the mechanisms of unconscious inferences (Friston & Frith, 2015).

1.3 From an instrumental to a constitutive role of interaction

Thus, despite these phenomenological concessions from the mainstream, what still requires more systematic scientific assessment is whether an intersubjective experience at the personal level may not actually be better explained by positing an equally intersubjective mechanism also at the unconscious subpersonal level , i.e. a mechanism that involves internal and relational activities from more than one person. In other words, it remains to be seen to what extent we can replace the mainstream restriction to subpersonal mechanisms consisting only of one individual's neural activity with a much broader perspective on subpersonal mechanisms consisting of all kinds of activity crisscrossing the brains, bodies and environment of two or more people.

This alternative possibility is being taken seriously by a growing number of researchers who adopt a more encompassing view of subpersonal mechanisms, for example in terms of interpersonal synergies (Chemero, 2016), autonomous dynamics of mutual interaction (De Jaegher, 2009) and inter-brain coordination (Di Paolo & De Jaegher, 2012). What these recent developments demonstrate is that, at least conceptually and methodologically, there is no longer any reason to limit our explanatory toolkits to fit methodological individualism. But are there actually any phenomena at the personal level that force us to move beyond this individualist restriction at the subpersonal level? That is, can we find empirical evidence that is consistent with genuine intersubjectivity such that there is co-constitution of one and the same shared moment of experience? Can we thereby ground our experience of being with others?

Fortunately, that finding this evidence is a promising possibility is suggested by an essential characteristic of second-person interaction, namely reciprocity (De Bruin, van Elk, & Newen, 2012). As an example let us consider in detail the reciprocity of emotional interaction, as illustrated in Figure 1.

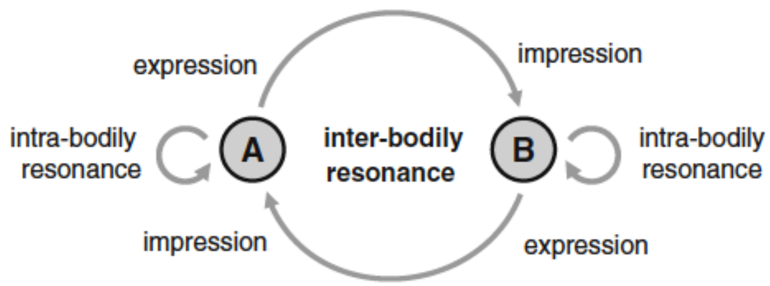

Figure 1. Illustration of pre-reflective intra- and inter-bodily affective resonance.

Reproduced from *Phenomenology and the Cognitive Sciences*, 11 (2), pp 205–235, The extended body: a case study in the neurophenomenology of social interaction, Tom Froese and Thomas Fuchs, doi:10.1007/s11097-012-9254-2 © Springer Science+Business Media B.V. 2012. With permission of Springer.

Following Froese and Fuchs (2012), let us assume that in this illustration person *A* undergoes an emotion, e.g. anger, which is manifested 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: person *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 intra-bodily resonance is an external expression at the same time: the anger becomes visible and is directly perceived as such by *A*'s partner *B*[1]. What is more, *A*'s expression will produce an impression in *B* by triggering corresponding or complementary bodily feelings. Thus, *A*'s sharpness of voice might induce in *B* an unpleasant tension, a tendency to withdraw, etc. Person *B* not only sees the anger immediately in *A*'s face and gesture, but also senses it with his body, through his own intra-bodily resonance. However, it does not stay like this, for the bodily reaction caused in *B* in turn becomes an expression perceived by *A* causing another impression; it will affect his bodily reaction, change his expression, however slightly, and so forth.

Froese and Fuchs (2012) argue that there is a circular interplay of expressions and reactions running in split seconds and constantly modifying each subject's bodily state, in a process that becomes highly autonomous and is not directly controlled by either of them. They have become parts of a dynamic sensorimotor and inter-affective system that connects both bodies by reciprocal movements and reactions, that means, in inter-bodily resonance. Both subjects experience a feeling of being connected with the other in a dynamic way that may be termed mutual incorporation (Fuchs & De Jaegher, 2009). Each lived body reaches out, as it were, to be extended by the other, dynamically forming an extended body. This is accompanied by a concrete, holistic impression of the interaction partner and a feeling for the atmosphere of the shared situation: you and me are having a tense argument that is out of control.

Note that, although this is an example of a confrontational situation rather than a cooperative one, according to the phenomenological tradition of intersubjectivity it is still a case of interdependent empathic understanding, where empathy is defined as a basic and intuitive experience of the embodied and expressive mind of the other (Zahavi, 2016). To be sure, this kind of second-person perspective is not always sufficient to constitute an integrated 'we' in which two subjects share a specific experience, such as two parents identifying with each other by sharing the joy of seeing their child's first steps. In contrast, in the example described by Froese and Fuchs, the interaction resulted in interdependent yet divergent and individual-specific emotions, e.g. anger in person *A* and a corresponding shock in person *B*. The emergence of a 'we', in which such 'I-you' interdependence is transformed into a new, qualitatively shared intersubjective perspective, may require further integration via collective concerns and values.

No appeal to theorizing and/or simulation is necessary in describing this kind of second-person interaction, at least not at the phenomenological level, although a subpersonal explanation adhering to representationalist-internalist mechanisms is probably always conceivable – for example by means of fully duplicating the other and our interaction as a theory/simulation in my head. On the other hand, why duplicate anything that is there in the situation already? In the general case of perception it has long been argued that internal models are not necessary if the world itself can play the required role, so it is worth considering whether we

[1] There is a growing literature about the conditions for the direct perception of emotions, which I will not enter into here (see, e.g., Stout, 2012).

could extend this idea to social scenarios. If it is possible to take advantage of the situational complexity of social interaction and thereby externalize some of the brain's social cognitive load, it is unlikely that evolution would have let such an opportunity go to waste (De Jaegher, et al., 2016)[2]. In other words, it is also possible that we could explain such interactive direct social understanding as being grounded in a pre-reflectively lived coupling, an inter-bodily reciprocity that has created a "mixture of myself and the other" (Merleau-Ponty, [1960] 1964, p. 155). This is an empirically verifiable hypothesis: if we go beyond the representationalist-internalist constraints of traditional cognitive science, might it be possible to find examples of such a mixture of self and other at the level of extended subpersonal mechanisms underlying shared experience?

The mirror neuron system might be the mainstream's most suitable starting point because it is activated no matter whether the action is executed by the self or perceived as something done by the other. Indeed, a time series of mirror neuron firing could therefore involve an indistinguishable mixture of self- and other-generated activity. But this self-other neutrality is also a problem because genuine second-person interaction actually depends on the maintenance of a distinction between you and me (Zahavi, 2016), and it has long been recognized that the mirror neuron system by itself is not sufficient for explaining self-other distinction (Jeannerod & Pacherie, 2004). And, more importantly, mirror neuron activation does not depend on reciprocal interaction; one individual's passive observation of another's action is sufficient as a trigger.

Instead we are looking for a mechanism that is reciprocally distributed across two subjects while they are still retaining their distinct individuality. Thus, a phenomenological analysis of intersubjectivity can lead to a very different conceptualization of how we come to know other minds, relegating reasoning by analogy or simulation to a derivative role dependent on direct perception, while instead emphasizing the more fundamental role of embodied co-determination of autonomous systems (Fuchs & De Jaegher, 2009). In the words of Merleau-Ponty:

> Between my consciousness and my body as I experience it, between this phenomenal body of mine and that of another as I see it from the outside, there exists an internal relation which causes the other to appear as the completion of the system. (Merleau-Ponty, [1945] 2002, p. 410)

On this view, one's engagement with another person is not something external to the process of perceiving them as such; instead it is constitutive of an irreducible second-person perspective of 'me and you' that reciprocally integrates the first-person perspectives of both subjects. Even more radically, it is plausible that

---

[2] A range of evolutionary robotics experiments, in which embodied agents are evolved to solve a task in an environment populated with other agents, supports this conjecture. Invariably, the evolutionary algorithm encounters solutions to the task whereby the mechanisms generating an agent's performance cannot be separated from its ongoing interaction with other agents (see, e.g., Di Paolo, Rohde, & Iizuka, 2008; Froese, Iizuka, & Ikegami, 2013).

such embodied interaction is our primary social capacity, developmentally and in adult life (Reddy, 2008), whereas the first- and third-person perspectives, as appealed to by simulation and theory versions of theory of mind, respectively, are the basis of derivative modes giving rise to more complex forms of social understanding (Fuchs, 2013; Merleau-Ponty, [1960] 1964; Gallagher, 2012). Roughly, on this view, it seems plausible that pre-reflective embodied interaction without much self-other distinction is the most fundamental social mechanism that enables more explicit forms of self-awareness, which gives rise to a more fully articulated second-order perspective, which then enables more advanced social interaction skills to be deployed, and so forth in a dialectical spiral. At some point this history of social interaction will enable the emergence of more individual forms of social cognition that have been the focus of theory of mind approaches, namely those involving detached reflection and pretense.

The upshot is that there is nothing standing in the way of the possibility of a cognitive science of genuine intersubjectivity. Moreover, if we do not specifically look for the involvement of the other when trying to account for social processes it is unlikely that we will find such involvement, thereby inadvertently turning methodological individualism into something like a self-fulfilling prophecy. It is time for a more comprehensive research program to be established in order to better elucidate the potential mechanisms of genuine intersubjectivity.

2. Toward a science of genuine intersubjectivity

The enactive approach to social cognition is in a promising position to meet the requirements (De Jaegher & Di Paolo, 2007; Froese & Di Paolo, 2011; Froese & Gallagher, 2012; Fuchs & De Jaegher, 2009). Its novel research program enables us to better integrate both empirical and phenomenological findings in order to determine to what extent embodied interaction can indeed give rise to a socially extended mind in the strong sense of genuine intersubjectivity. Importantly, it argues that social interaction dynamics can be constitutive of social cognition (De Jaegher, Di Paolo, & Gallagher, 2010; Di Paolo & De Jaegher, 2012), which is in line with phenomenological analyses of intersubjectivity, for example of how the second-person perspective is constituted by the reciprocal interlocking of two first-person perspectives (León & Zahavi, 2016). The enactive approach, in agreement with the phenomenological tradition, therefore marks a clear break with theory of mind approaches. Yet it can also be seen as a development arising from within the history of cognitive science, as illustrated in Figure 2.

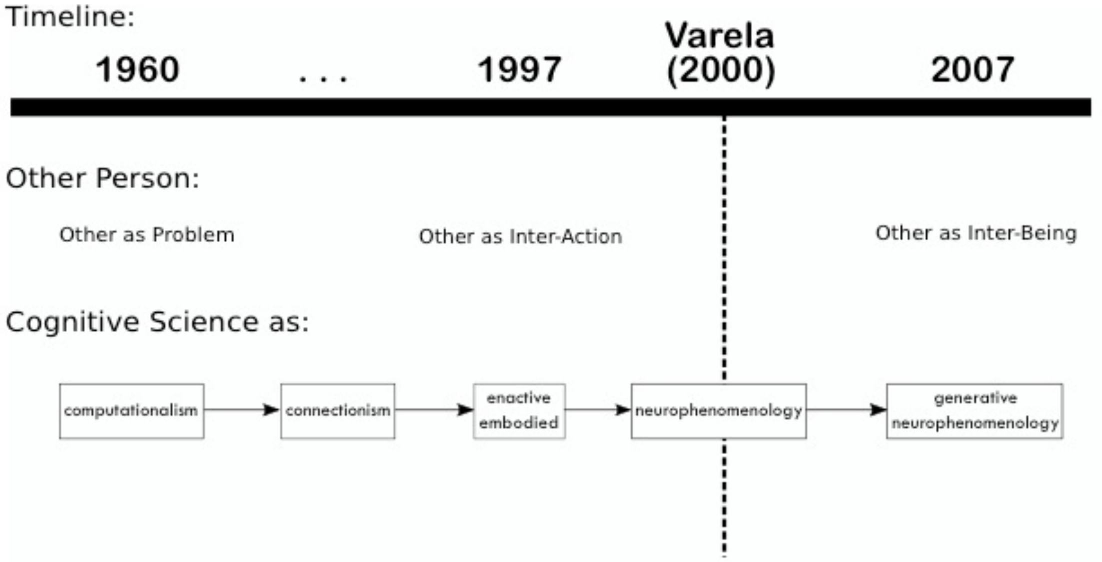

Figure 2. Steps to a science of inter-being. According to Varela (2000), the successive stages in the history of cognitive science can be seen as corresponding to a transformation in the scientific understanding of the other. He projected that in the 2000s there will arise a cognitive science of inter-being, what we have called the enactive approach to genuine intersubjectivity.

As we have already seen, for classical cognitive science, here represented by the paradigms of computationalism and connectionism, the other's mind is mostly a problem: the self is an isolated agent who must figure out how to survive in a world populated by potentially mindless bodies. For embodied cognitive science and static neurophenomenology, i.e. approaches that acknowledge the role of embodied interaction and direct perception without changing their underlying presuppositions about the limits of constitution, the other is mainly conceived of in terms of an affordance for interaction; the appearance of the other is a given fact but external links must still be built between two independent minds. At this intermediate stage, which is the mainstream's state of the art (Wiltshire, et al., 2015), there is still no room for Merleau-Ponty's radical concept of an internal relation that allows the other to be the completion of a larger, integrated system.

Yet Varela already foresaw the seeds of a future stage of cognitive science in which genuine intersubjectivity, what he calls inter-being, i.e. an interaction in which we co-constitute a shared awareness in contrast to an interaction that remains external input to independent individuals, may be thematized without contradiction: "the other and I are common ground, a joint tissue which is tangibly present in empathy and affect, which offer a possible level of analysis if we avail ourselves of the means to do so" (Varela, 2000, p. 87). However, while Varela was optimistic about this transformation of social cognition research, he was also cautious about the speed of acceptance of the alternative view given how ingrained methodological individualism and skepticism of phenomenology are in the modern worldview.

> These are heavily inertial assumptions that will move as slowly as continents. The natural attitude of the scientist and the public today is to see the mind as a distinct, brain-encased self. Breaking that illusion from within science seems, today, not a complete impossibility – some cracks are opening for a science of interbeing. (Varela, 2000, p. 87)

Indeed, such cracks have been appearing since Varela published his projected timeline, with contributions picking up speed and recognition since a decade ago.

2.1 Methodology

We can group these contributions into three major pillars that together form the overall enactive research program: phenomenology, theory, and experiment, as illustrated in Figure 3. As we have seen, much progress has already been made in terms of the first two pillars. Classical and contemporary phenomenological philosophy provide descriptions of intersubjectivity, including empathy, direct social perception, second-person perspective, intercorporeity, etc. (Gallagher & Zahavi, 2008). The theory has been developed by the enactive approach to social cognition, which emphasizes the constitutive role of social interaction in terms of a variety of concepts, such as self-other co-determination (Thompson, 2001), participatory sense-making (De Jaegher & Di Paolo, 2007), mutual incorporation (Fuchs & De Jaegher, 2009), self-other contingencies (McGann & De Jaegher, 2009), and the extended body (Froese & Fuchs, 2012). This is not the place to compare and contrast all of these different concepts. What they share in common is the aim to highlight that the embodied mind is intersubjectively constituted at its most fundamental levels.

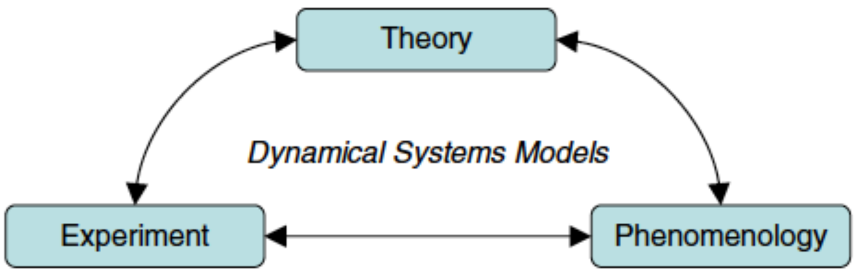

Figure 3. The three pillars of the enactive research program. The overall aim is to integrate theory, experiment, and phenomenology into one coherent framework by mutually informing and constraining interdependencies. Ideally, the three pillars should be formally integrated by means of dynamical systems theory. When the experiments are conducted in the field of neuroscience, this research program is known as neurophenomenology (e.g., Varela, 1999).

However, there continues to be a lamentable disconnect between this growth in phenomenological and theoretical research on the one hand and the relative lack of experiments on genuine intersubjectivity on the other. To be fair, there are some compelling psychological studies on the potentially constitutive role of real-time embodied interaction in social cognition (Auvray & Rohde, 2012;

Dumas, Kelso, & Nadel, 2014; De Jaegher, et al., 2010). And there is a growing sophistication of hyperscanning methods in social cognitive neuroscience, which have revealed various ways in which brains mutually influence each other during online social interaction (Dumas, Laroche, & Lehmann, 2014; Chatel-Goldman, Schwartz, Jutten, & Congedo, 2013; Schilbach et al., 2013; Hari, Henriksson, Malinen, & Parkkonen, 2015). But so far there exists not a single study aimed at specifically verifying if genuine intersubjectivity is possible by integrating all of the methodological elements shown in Figure 3.

One important factor is that many of the studies on social interaction whose findings might be interpreted from the perspective of genuine intersubjectivity are conceived and analyzed from within the strict confines of methodological individualism (Chatel-Goldman, et al., 2013). Accordingly, no evidence of social interaction could ever amount to evidence of Varela's inter-being if the latter is not even conceivable in principle. Another important factor is that none of these studies included a systematic assessment of the experience that participants undergo during their interaction. Here the enactive approach collides with yet another deep-seated prejudice found in experimental cognitive science, namely that subjective reports are not trusted as reliable data. This premise is also slowly changing, partly because of the concerted push by phenomenological and enactive approaches for lived experience to be taken seriously (Froese, Gould, & Barrett, 2011), and partly because the burgeoning field of consciousness science is in need of more reliable methods of measurement (Sandberg, Timmermans, Overgaard, & Cleeremans, 2010). In general, attitudes toward consciousness are changing, even if important practical difficulties remain.

For the enactive approach to social cognition the more fundamental of these two problematic factors is the continued insistence on methodological individualism as the gold standard of acceptability in the sciences of the mind. For as we have seen, even if it has become more accepted that we have the experience of directly perceiving other minded beings and sharing experiences with them, this insight by itself is not sufficiently compelling for most cognitive scientists to expand their traditionally internalist scope of subpersonal explanations so as to consider social interaction dynamics in themselves as a part of the necessary causal or even constitutive mechanisms. To change this conservative attitude it would help if it could be demonstrated that in principle there is nothing mysterious about the theoretical shift from treating an interaction as an external affordance for *independent* individuals to conceiving their interaction as an opportunity for the establishment of an internal relation between *interdependent* individuals.

2.2 Agent-based modeling as proof of concept

Due to this conceptual requirement of demonstrating an in-principle possibility, dynamical systems theory has been playing a somewhat different role in the enactive approach to social cognition, compared to how Varela (1999) had envisioned its place in his neurophenomenology of time consciousness. For him a key advantage of explanations in terms of dynamical system theory, in contrast to the representational terminology of the cognitive unconscious, is its neutrality with respect to scale and domain. The theory can account for and integrate activity in brain, body and environment, including interactions with others, by abstractly modeling the relevant dynamics (Kelso, 1995). Moreover, it therefore does not matter whether the phenomena to be explained are subjective or objective; activity in both domains is unfolding in time and can therefore be integrated into one explanatory framework.

Finally, and here we briefly return to the mainstream's problematic assumption of the cognitive unconscious, using dynamical systems theory to explain what is going on at the subpersonal level has another double advantage: 1) it makes these accounts formally continuous with the standard mathematics of the rest of the natural sciences, in contrast to the fundamental break introduced by the mainstream's appeal to homuncular or representational mechanisms, and 2) it thereby helps to refocus the study of subjectivity to where it belongs – in the concrete lifeworld of the personal level – rather than projecting its abstractions into the hypothetical domain of the cognitive unconscious.

These advantages are retained in the study of genuine intersubjectivity, but the aim of formally integrating concrete subjective and objective descriptions into one dynamical account faces considerable difficulties at the moment. Currently we do not have sufficiently fine-grained descriptions of the real-time interaction dynamics in its objective and especially its subjective dimensions. Instead we make use of another key advantage of the dynamical approach, namely that it lends itself to creating and analyzing minimal agent-based models of cognition (Beer, 2000). These simulation models can also be related to the three pillars of the enactive research program (Figure 3), e.g. by helping to formalize theoretical claims and phenomenological insights and by providing fresh perspectives on psychological and neuroscientific studies (Froese & Gallagher, 2012; Rohde, 2010).

One of the most basic insights of the dynamical approach is that all behavior is a relational phenomenon belonging to the whole brain-body-environment system (Beer, 2000). Accordingly, all behavior is extended and must be explained in an appropriately distributed manner. Social behavior is no exception, as confirmed by a series of modeling studies, which lead us to conceive of social interaction as a relational property of one integrated brain-body-environment-body-brain system (Froese, Iizuka, et al., 2013). This idea is illustrated in Figure 4.

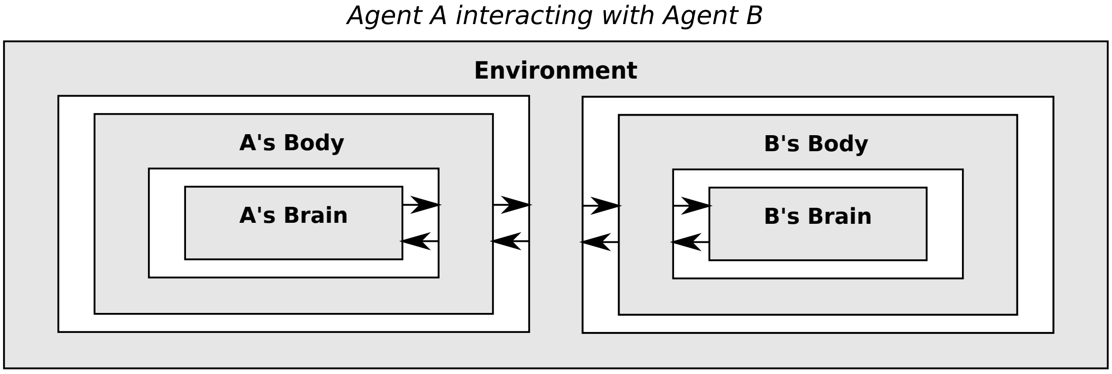

Figure 4. A dynamical perspective on the interaction between two situated, embodied agents. An agent's nervous system (abbreviated as "brain"), body, and environment are conceptualized as nonlinear

dynamical systems that are parametrically coupled, thus forming an irreducible whole. When agent *A* is interacting with agent *B*, their mutual coupling constitutes a whole brain–body–environment–body–brain system.

What this dynamical perspective to social interaction implies is that changes to neural, bodily, environmental, or intersubjective conditions can lead to changes at level of the system as a whole, which in turn changes the conditions of the components. In this way social interaction is constitutive of social cognition: an individual agent's social behavior depends on the coupling of all the subsystems, including the other agent, and cannot be attributed to any one component in isolation. Importantly, what this means is that the brain is not an isolated black box that must try to represent what is happening in the outside world in order to predict what will happen next and to infer what it ought to do in response (Clark, 2013; Dennett, 1987). Rather, the brain is an important, even essential, part of the whole intersubjective situation and is thereby directly participating in the social interaction (Gallagher, Hutto, Slaby, & Cole, 2013). We can illustrate this implication in terms of a minimalist agent-based model. This model suffers from obvious practical limitations, such as reducing an agent's situatedness and embodiment to nothing but being a moving object in an empty space, as well as from in principle limitations, especially the lack of an experiential dimension. Nevertheless, it is also precisely because the model's minimalism that we stand a chance to understand its emergent behaviors in some detail.

Froese and Fuchs (2012) describe a computer simulation in which two agents are situated in an open-ended 1D space along which they can move left- and rightwards. One agent is oriented upward and the other downward. The agents are equipped with a touch sensor, which becomes active when the bodies of the agents are overlapping and remains turned off otherwise. The artificial brain of each agent is a standard continuous-time recurrent neural network that consists of three fully interconnected neurons. Activity of all the neurons is modulated by the status of the touch sensor, and two dedicated motor neurons modulate the velocity of movement, thereby forming a closed sensorimotor loop.

The parameters of the neural network (weights, time constants, biases) and of the way in which it is coupled with the body and environment (sensor and motor gains) are automatically shaped by an evolutionary algorithm, which selects pairs of structurally identical agents according to how well they coordinate their behavior. The task consists of finding each other, negotiating a shared direction of movement, and jointly moving toward that direction while repeatedly making contact. Once adequate social interaction has evolved it is possible to analyze the state space of an isolated neural network, as shown in Figure 5.

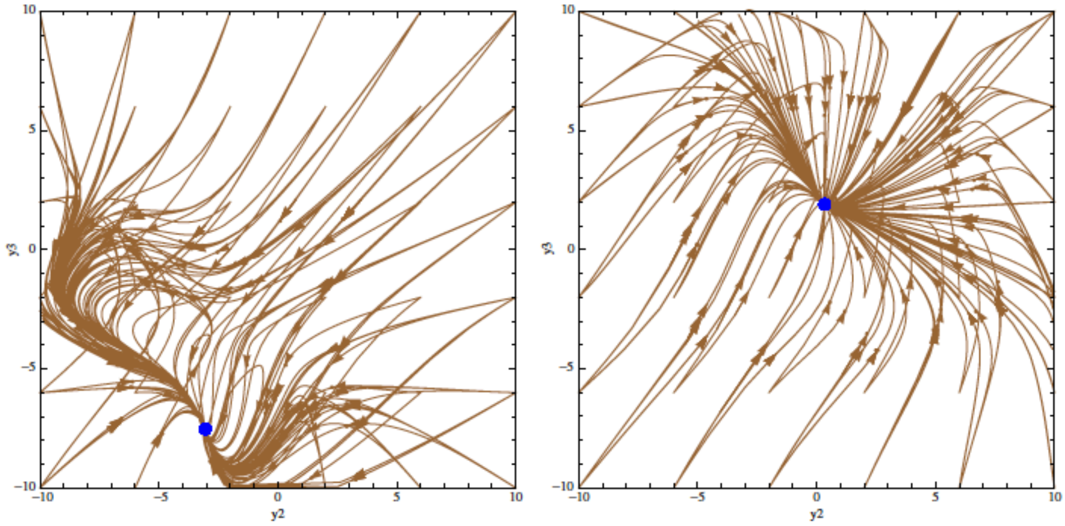

Figure 5. Representative flow structure of a section of the activation space of an agent's continuous-time recurrent neural network. Notice how the structure of this system changes depending on whether the contact sensor is fixed to off (left) or on (right).

To facilitate the analysis we focus only on the activation vector of a subsystem consisting of the two motor neurons y_2 and y_3, whose outputs modulate left- and rightward components of overall velocity, respectively. To isolate the agent from its environment we fix its sensor to be continuously off or on. In both conditions flows of activation eventually converge on a single stable fixed-point attractor. However, the position of this equilibrium point and the structure of its basin of attraction differ depending on the state of the sensor. In other words, for this neural network making contact does not merely serve as an external input to be processed by a fixed subpersonal architecture. Rather, contact has the effect of reorganizing the dynamical structure of the neural network's motor subsystem, such that the contact itself cannot be considered as separate from the way in which the agent's behavior is realized But this point holds equally for all forms of sensorimotor interaction, not just for the special case when the neural activation is generated via coordinated interaction with another agent (see also Chemero, 2016). What happens with the state of the motor subsystem during conditions of social interaction is qualitatively different, as can be seen in **Figure 6.**

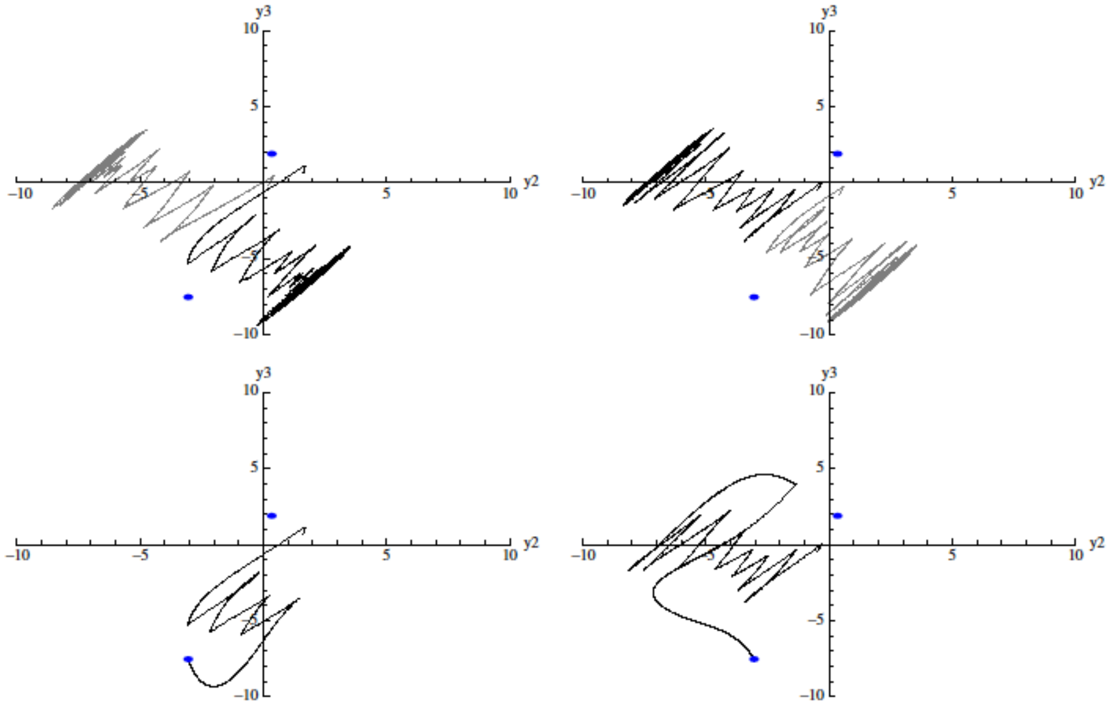

Figure 6. Trajectories in neural activation space for representative trials of the 'live' interaction condition (top row) and 'playback' condition (bottom row), while heading leftward (left column) or rightward (right column). Activations of agent A (black line) and agent B (gray line) always start near the origin and then progress outwards with time. The locations of the normally separately existing equilibrium points, i.e. one for each sensor status, are superimposed simultaneously for ease of reference (solid dots).

Reproduced from *Phenomenology and the Cognitive Sciences*, 11 (2), pp 205–235, The extended body: a case study in the neurophenomenology of social interaction, Tom Froese and Thomas Fuchs, doi:10.1007/s11097-012-9254-2 © Springer Science+Business Media B.V. 2012. With permission of Springer.

It turns out that the motor subsystem behaves very differently when both agents co-regulate their movements (top row of Figure 6), compared to when an agent is isolated by fixing its input (**Figure 5**). During reciprocal interaction the agents' activations are maintained as one of two far-from-equilibrium transients, which can take place in two distinct regions of state space depending on whether the agents are jointly moving leftward or rightward. These self-organized transients of neural activity are the cause and effect of the co-constitution of a robust yet flexible coordination of behaviors. In other words, reciprocal sensorimotor interaction between the two agents spontaneously transforms the dynamical organization of their motor subsystems, permitting them to jointly generate more complex patterns of behavior than would be possible for either agent in isolation.

Despite the model's simplicity (or rather thanks to it), this finding might be a good starting point for formalizing Merleau-Ponty's claim that our perceptions of others "arouse in us a reorganization of motor conduct, without our already having learned the gestures in question" ([1960] 1964, p. 145). Moreover, it also provides us with a clue to formalizing his claim that when perceiving others "there exists an internal relation which causes the other to appear as the completion of the system" ([1945] 2002, p. 410). This is because the maintenance of the coordinated behavior depends on the active participation of the other agent. When we prevent the possibility of a reciprocal modulation of neural dynamics, for example when we re-initialize a trial but replace the other agent with a playback of its movements that were recorded during a previous run of that trial, the complex style of transient neural dynamics can no longer be maintained by the remaining active agent (**Figure 6**, bottom row).

This minimal agent-based model therefore serves as a formal proof of concept that an extended subpersonal mechanism that is compatible with the concept of genuine intersubjectivity, i.e. such as when we reciprocally participate in the interactive realization of each other's socially contingent behavior, is possible in principle. We are therefore in a suitable position to verify that this possibility can also be empirically confirmed in terms of actual psychological experiments of social interaction, in particular those that also take into account the experiential dimension of the participants.

2.3 Perceptual crossing experiments

Dynamical systems theory casts doubt on the foundational assumptions of mainstream cognitive science by questioning the necessity of its internalist-representationalist framework. In particular, conceiving of cognitive systems as nonlinearly coupled organism-environment systems "removes the pressure to treat one portion of the system as *representing* other portions of the system" (Silberstein & Chemero, 2012, p. 40). In addition, it leads us to expect that "the processes crucial for consciousness cut across brain-body-world divisions, rather than being brain-bound neural events" (Thompson & Varela, 2001, p. 418; but see Clark, 2009). Moreover, given the proof of concept provided by the agent-based model (and many others similar to it), it suggests that there is nothing standing in the way of generalizing these implications to a multi-agent system nonlinearly integrated via social interaction, that is, in which the world of each agent includes other agents. The next step is therefore to empirically verify that the processes crucial for experiences involving others can cut across self-other divisions, as required by genuine intersubjectivity.

Methodologically, this means that we are looking for an experimental paradigm that enables us to systematically investigate the extent to which second-person interaction dynamics allows subjects to participate in each other's experience, while still retaining the tractability of the minimalist agent-based models. One suitable way is to mediate participants' sensorimotor loops via minimal human-computer interfaces that are designed to distill the full complexity of human interaction capacities to its essential features (Rohde, 2010). For the study of embodied social interaction, a design that fulfills these requirements is the so-called perceptual crossing paradigm (Auvray & Rohde, 2012), which was first popularized by Auvray, Lenay, and Stewart (2009).

Auvray, Lenay, and Stewart (2009) set out to investigate whether recognition of another's presence can be brought about on the basis of embodied interaction in a minimalist virtual environment. No other interaction between the participants is possible other than binary tactile feedback. The basic experimental setup, with minor modifications introduced by Froese, Iizuka and Ikegami (2014), is shown in Figure 7. Each participant's interface consists of two parts: a trackball that controls the linear displacement of their virtual avatar, and a hand-held haptic feedback device that vibrates at constant frequency for as long as the avatar overlaps with another virtual object and remains off otherwise. Participants are separated by a wall and wear noise-canceling headphones. Three small lights on each desk signal the start, halftime (30 s), and completion of each 1-min trial.

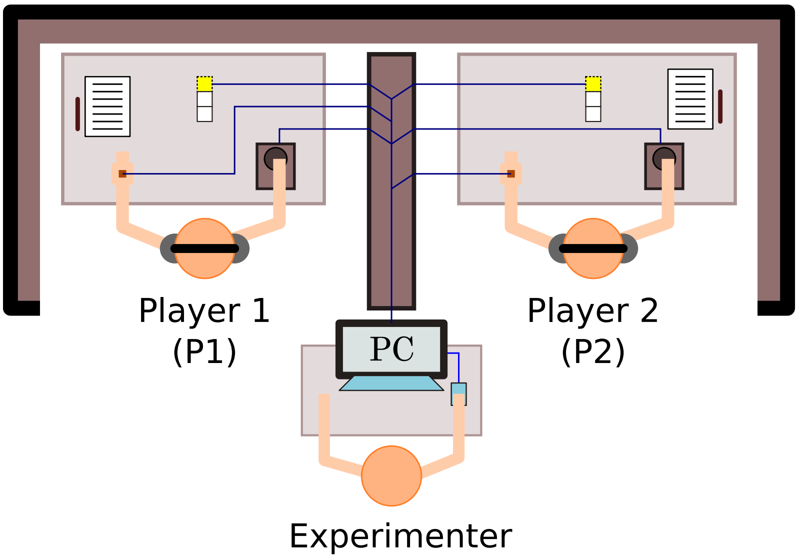

Figure 7. Experimental setup of a perceptual crossing experiment. The two players can only engage with each other via a human-computer interface that reduces their scope for embodied interaction to a bare minimum of translational movement and binary tactile sensation.

Players are virtually embodied as "avatars" in an invisible 1D space that wraps around, as illustrated in Figure 8. A player can encounter three types of object that have equal sizes and give rise to the same constant haptic feedback upon contact: a static object and two moving objects, one of which is the other player's virtual avatar and the other is an irresponsive object that directly copies the other's movements. Participants are informed about these three types of object, although they are not told that the irresponsive mobile object is an immediate copy of the other's movements. The point of this 'shadow' object is to determine if players are sensitive to the reciprocity of interaction as such, i.e. to its social contingency, rather than to just to the presence of external motion alone. The task given to each player is to click whenever they judged themselves to be interacting with the other player's avatar. No explicit feedback regarding the correctness of a participant's click is provided during the experiment and a click is unperceivable to the other participant.

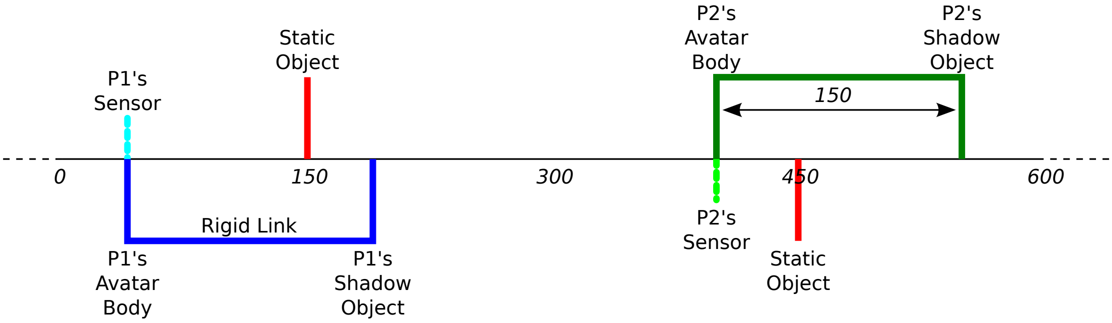

Figure 8. Virtual environment of a perceptual crossing experiment. Each avatar consists of a binary contact sensor and a body object. Unbeknownst to the players a 'shadow' object is attached to each avatar body at a fixed distance of 150 units. There are also two static objects, although only one is noticeable to each player. All objects are 4 units long and can therefore only be distinguished interactively in terms of their difference affordances for engagement.

Auvray, Lenay, and Stewart (2009) discovered that even under these minimalist and ambiguous conditions, most clicks happened during interaction with the other participant. The study was therefore a success. However, a closer look at the results challenged the most straightforward explanation, i.e. that participants clicked correctly because they were sensitive to social contingency. Participants were indeed less likely to click following tactile stimulation by a static object, but there was no significant difference between the probability of a click following stimulation by the other's avatar or by their shadow object. In other words, this suggests that individuals could not properly distinguish between the two moving objects from their own perspective.

So how is it possible that the experiment was a success? It turns out that while participants were not more likely to respond to contact with the other player, it was nevertheless more likely that they were in fact in contact with them than with the shadow object – thereby guaranteeing that most clicks following a stimulation by a moving object were in fact correct. And the reason why they tended to spend more time with each other was an emergent property of the situation: mutual interaction was relatively more stable than interaction with the other's shadow, given that in the latter situation the other player would still be searching for the partner and thereby dragging their shadow along with them. A one-sided interaction is more unstable and difficult to maintain, as we already saw from the spontaneous breakdown of interaction during the playback condition of the agent-based model.

This first perceptual crossing study was welcomed by many enactivists as an experimental demonstration of the claim that social interaction can constitute social cognition, at least under some conditions (De Jaegher, et al., 2010). Yet this radical interpretation of the results has received extensive criticisms and may not be tenable in its original form, at least not if we take cognition to be a form of sense-making, that is, as a process of which an essential component is its being meaningful for the agent. Importantly, the experiment provided no evidence that social cognition was actually shaped by the social interaction. To the contrary, participants were equally likely to click no matter whether the moving object that had caused the stimulation afforded a reciprocally co-regulated interaction or not. The relatively enhanced stability of the reciprocal interaction dynamics may have permitted the external self-organization of the objective correctness of the clicks, but that notable achievement is arguably not sufficient to talk about constitution of cognition (Overgaard & Michael, 2015). For the social interaction to be essentially constitutive, rather than just contextually enabling, we expect something to be qualitatively different for the participants during reciprocal interaction that marks it as a social situation, whether by transformation of the pre-reflectively felt meaning, perceptual experience, or judgment of the situation, which should have been reflected in their clicking behavior (for an alternative interpretation of constitution, see De Jaegher, et al., 2016).

Moreover, the null result of a lack of co-constitution of a genuine second-person perspective should actually be expected according to the enactive approach. As Froese and Di Paolo (2011) have argued, just as sense-making in the individual case requires the regulation of interaction rather than just interaction per se, sense-making in the intersubjective case requires the co-regulation of reciprocal interaction rather than just individually regulated interaction that coincidentally happened to be with another subject. Co-regulation signifies that the conditions of normativity are distributed across the agents. For example, the success of a handshake cannot be reduced to one person alone; the other person has to make a corresponding gesture. And the original experiment's task of agency detection does not require such co-regulation: we can imagine an extreme case in which a participant does not move at all and simply clicks when receiving consecutive stimulations, which is indicative of the searching activity of the other player.

Froese, Iizuka, and Ikegami (2014) set out to experimentally verify Froese and Di Paolo's argument by changing the instructions given to the players at the start of the perceptual crossing experiment, and by collecting subjective reports about the experience of the interaction preceding each click. Crucially, they told pairs of participants that they formed a team and that they needed to help each other inside the virtual environment to recognize each other's presence because their combined clicking performance would be ranked against other teams. The aim of these pro-social instructions was to encourage the formation of a 'we'.

This variation of the perceptual crossing paradigm resulted in several interesting findings. First, in contrast to the study by Auvray et al. (2009), participants were significantly more likely to click on the other's avatar than on the other's shadow, which provides compelling empirical support for the claim that social interaction can be constitutive of social cognition, even when adopting stricter criteria such as that the interaction process must also give rise to qualitative changes at the personal level. Second, in trials when participants jointly clicked correctly on each other, these clicks were more likely to lead to reports of a significantly clearer experience of the presence of the other subject, as measured by ratings using an adapted perceptual awareness scale (Sandberg, et al., 2010). Third, clicks in trials with joint success and associated with clear social perception were preceded by increased co-regulation of behavior, as measured by the extent of turn-taking between players. Turn-taking is interesting because its normativity is structured like a hand-shake: you cannot successfully turn-take by yourself. Fourth, the delays between jointly successful clicks tended to be small, with 15% occurring within two seconds of each other. This is suggestively close to the '1s' time scale that Varela (1999) took to be indicative of the time taken by large-scale neural integration during the emergence of a coherent cognitive act, at least for the case of one individual. Could a similar mechanism of neural integration be at work when the process involves more than one brain? This is an intriguing hypothesis. However, to further support this possibility of large-scale inter-brain integration, we would need to conduct a perceptual crossing study that includes a double-EEG analysis of the interacting players.

3. Conclusions

Overall, enactive theory, models, and experiments converge on the claim that co-regulated social interaction constitutes genuine intersubjectivity, such that it is indeed possible for us to mutually participate in the generation of our intentions, emotions, and experiences more generally. This is a desirable outcome because it means that cognitive science can be made consistent with the folk psychological intuition that the presence of others makes a difference to how we behave, think, and live. Others transform how the world shows up for us.

In contrast, according to the internalist starting point adopted by mainstream social cognition research, all aspects of sociality must ultimately be consistent with an overall ontology and metaphysics in which isolated individuals (or even just their brains) are the sole repositories of intentionality, whether individual or collective. If the mainstream were to allow even the mere possibility of forms of cognition whose underlying mechanisms are not reducible to what is going on within isolated individuals, then methodological individualism would have to be replaced with a framework that is at least potentially capable of inter-individual explanations. This would have far-ranging repercussions for how we understand ourselves as well as our place in this world.

Searle is to be commended for making explicit the constraints imposed on our understanding by mainstream cognitive science:

> (1) It must be consistent with the fact that society consists of nothing but individuals. Since society consists entirely of individuals, there cannot be a group mind or group consciousness. All consciousness is in individual minds, in individual brains.
> (2) It must be consistent with the fact that the structure of any individual's intentionality has to be independent of the fact of whether or not he is getting things right […] One way to put this constraint is to say that the account must be consistent with the fact that all intentionality, whether collective or individual, could be had by a brain in a vat or by a set of brains in vats. (Searle, 1990, pp. 406-7)

Searle recognizes that these constraints resemble methodological individualism and methodological solipsism, respectively, although he prefers to construe them "as just commonsensical, pre-theoretical requirements." Yet it should be evident by now that these requirements are neither commonsensical nor pre-theoretical; they are deeply rooted in the internalist commitments of mainstream tradition of cognitive science To be fair, many proponents of this tradition probably pursue their work without considering the possibility of solipsism. They prefer to treat others as real subjects in their own right, albeit with the caveat that they are still forced to claim that their existence is epiphenomenal from the perspective of scientific accounts of social cognition. To be able to scientifically recognize that the presence of others does make a difference in a meaningful way we must first overcome the constraints identified by Searle.

More importantly, as argued in this chapter, it is possible to break free from these constraints while staying within the remit of natural science. Rejecting them does not lead to accepting the existence of a mysterious group mind or consciousness (Boden, 2006). Rather, it allows us to acknowledge and properly investigate the role of social interaction dynamics in constituting individual and intersubjective capacities. Once this turn toward the constitutive role of embodiment and interaction is undertaken, the brain in a vat thought experiment looses its intuitive force (Thompson & Cosmelli, 2011). And accepting that we sometimes get things wrong does not entail that we are never right about experiencing ourselves to be in direct relation with others and the world around us (Beaton, 2013).

The upshot is that the enactive approach is bringing about a shift in mainstream ontology and metaphysics toward a worldview that is more consistent with the phenomenological approach to social reality: reciprocal interaction with other people plays a constitutive role in implicitly and explicitly shaping our lives and minds. We genuinely have relationships with others; we are not independent but interdependent. These insights return cognitive science closer to actual human experience, and thereby allow us to better appreciate and address the growing challenges of living in an increasingly interconnected world.

Acknowledgements

Tom Froese thanks Guillaume Dumas and Jonas Chatel-Goldman for the many fruitful discussions. Thanks also to Ezequiel Di Paolo for his attentive comments on an earlier version of this manuscript. Detailed feedback provided by an anonymous reviewer helped to further fine-tune the arguments. This work was supported by SEP-CONACyT project #221341 CB-2013-01 ("Análisis de redes complejas en sistemas biológicos y sociales").

References

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

Auvray, M., & Rohde, M. (2012). Perceptual crossing: The simplest online paradigm. *Frontiers in Human Neuroscience, 6*(181). doi: 10.3389/fnhum.2012.00181

Beaton, M. (2013). Phenomenology and embodied action. *Constructivist Foundations, 8*(3), 298-313.

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

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

Bohl, V., & Gangopadhyay, N. (2014). Theory of mind and the unobservability of other minds. *Philosophical Explorations, 17*(2), 203-222.

Carruthers, P. (2015). Perceiving mental states. *Consciousness and Cognition, 36*, 498-507.

Chatel-Goldman, J., Schwartz, J.-L., Jutten, C., & Congedo, M. (2013). Non-local mind from the perspective of social cognition. *Frontiers in Human Neuroscience, 7*(107). doi: 10.3389/fnhum.2013.00107

Chemero, A. (2016). Sensorimotor empathy. *Journal of Consciousness Studies, 23*(5-6), 138-152.

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

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181-204.

De Bruin, L., van Elk, M., & Newen, A. (2012). Reconceptualizing second-person interaction. *Frontiers in Human Neuroscience, 6*(151). doi: 10.3389/fnhum.2012.00151

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

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

De Jaegher, H., Di Paolo, E. A., & Adolphs, R. (2016). What does the interactive brain hypothesis mean for social neuroscience? A dialogue. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*(20150379). doi: 10.1098/rstb.2015.0379

De Jaegher, H., Di Paolo, E. A., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences, 14*(10), 441-447.

Dennett, D. C. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.

Di Paolo, E. A., & De Jaegher, H. (2012). The interactive brain hypothesis. *Frontiers in Human Neuroscience, 6*(163), doi: 10.3389/fnhum.2012.00163.

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.

Dumas, G., Kelso, J. A. S., & Nadel, J. (2014). Tackling the social cognition paradox through multi-scale approaches. *Frontiers in Psychology, 5*(882). doi: 10.3389/fpsyg.2014.00882

Dumas, G., Laroche, J., & Lehmann, A. (2014). Your body, my body, our coupling moves our bodies. *Frontiers in Human Neuroscience, 8*(1004). doi: 10.3389/fnhum.2014.01004

Friston, K. J., & Frith, C. D. (2015). A duet for one. *Consciousness and Cognition, 36*, 390-405.

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

Froese, T., & Fuchs, T. (2012). The extended body: A case study in the neurophenomenology of social interaction. *Phenomenology and the Cognitive Sciences, 11*(2), 205-235.

Froese, T., & Gallagher, S. (2012). Getting interaction theory (IT) together: Integrating developmental, phenomenological, enactive, and dynamical approaches to social interaction. *Interaction Studies, 13*(3), 436-468.

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.

Froese, T., Iizuka, H., & Ikegami, T. (2013). From synthetic modeling of social interaction to dynamic theories of brain-body-environment-body-brain systems. *Behavioral and Brain Sciences, 36*(4), 420-421.

Froese, T., Iizuka, H., & Ikegami, T. (2014). Embodied social interaction constitutes social cognition in pairs of humans: A minimalist virtual reality experiment. *Scientific Reports, 4*(3672). doi: 10.1038/srep03672

Froese, T., & Leavens, D. A. (2014). The direct perception hypothesis: Perceiving the intention of another's action hinders its precise imitation. *Frontiers in Psychology, 5*(65). doi: 10.3389/fpsyg.2014.00065

Froese, T., Stanghellini, G., & Bertelli, M. O. (2013). Is it normal to be a principal mindreader? Revising theories of social cognition on the basis of schizophrenia and high functioning autism-spectrum disorders. *Research in Developmental Disabilities, 34*(5), 1376-1387.

Fuchs, T. (1996). Leibliche Kommunikation und ihre Störungen. *Zeitschrift für klinische Psychologie, Psychopathologie und Psychotherapie, 44*, 415-428.

Fuchs, T. (2013). The phenomenology and development of social perspectives. *Phenomenology and the Cognitive Sciences, 12*, 655-683.

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

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

Gallagher, S. (2012). In defense of phenomenological approaches to social cognition: Interacting with the critics. *Review of Philosophy and Psychology, 3*(2), 187-212.

Gallagher, S., Hutto, D. D., Slaby, J., & Cole, J. (2013). The brain as part of an enactive system. *Behavioral and Brain Sciences, 36*(4), 421-422.

Gallagher, S., & Zahavi, D. (2008). *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science*. London, UK: Routledge.

Gallese, V. (2005). Embodied simulation: From neurons to phenomenal experience. *Phenomenology and the Cognitive Sciences, 4*(1), 23-48.

Gallotti, M., & Frith, C. D. (2013). Social cognition in the we-mode. *Trends in Cognitive Sciences, 17*(4), 160-165.

Hari, R., Henriksson, L., Malinen, S., & Parkkonen, L. (2015). Centrality of social interaction in human brain function. *Neuron, 88*(1), 181-193.

Hyslop, A. (2016). Other minds. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2016 ed.).

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

Kelso, J. A. S. (1995). *Dynamic Patterns: The Self-Organization of Brain and Behavior*. Cambridge, MA: The MIT Press.

Kiverstein, J. (2011). Social understanding without mentalizing. *Philosophical Topics, 39*(1), 41-65.

Krueger, J. (2012). Seeing mind in action. *Phenomenology and the Cognitive Sciences, 11*(2), 149-173.

León, F., & Zahavi, D. (2016). Phenomenology of experiential sharing: The contribution of Schutz and Walther. In A. Salice & H. B. Schmid (Eds.), *The Phenomenological Approach to Social Reality* (pp. 219-234). Switzerland: Springer.

McGann, M., & De Jaegher, H. (2009). Self-other contingencies: Enacting social perception. *Phenomenology and the Cognitive Sciences, 8*(4), 417-437.

Merleau-Ponty, M. ([1945] 2002). *Phenomenology of Perception* (C. Smith, Trans.). Oxon, UK: Routledge.

Merleau-Ponty, M. ([1960] 1964). The child's relations with others. In J. M. Edie (Ed.), *The Primacy of Perception* (pp. 96-158). Evanston, IL: Northwestern University Press.

Michael, J., & De Bruin, L. (2015). How direct is social perception? *Consciousness and Cognition, 36*, 373-375.

Overgaard, S. (2012). Other people. In D. Zahavi (Ed.), *The Oxford Handbook of Contemporary Phenomenology* (pp. 460-479). Oxford: Oxford University Press.

Overgaard, S., & Michael, J. (2015). The interactive turn in social cognition research: A critique. *Philosophical Psychology, 28*(2), 160-183. doi: 10.1080/09515089.2013.827109

Ratcliffe, M. (2007). *Rethinking Commonsense Psychology: A Critique of Folk Psychology, Theory of Mind and Simulation*. New York, NY: Palgrave Macmillan.

Reddy, V. (2008). *How Infants Know Minds*. Cambridge, MA: Harvard University Press.

Rohde, M. (2010). *Enaction, Embodiment, Evolutionary Robotics: Simulation Models for a Post-Cognitivist Science of Mind*. Amsterdam, Netherlands: Atlantis Press.

Sandberg, K., Timmermans, B., Overgaard, M., & Cleeremans, A. (2010). Measuring consciousness: is one measure better than the other? *Consciousness and Cognition, 19*(4), 1069-1078.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences, 36*(4), 393-462.

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

Silberstein, M., & Chemero, A. (2012). Complexity and extended phenomenological-cognitive systems. *Topics in Cognitive Science, 4*(1), 35-50.

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

Stout, R. (2012). What someone's behavior must be like if we are to be aware of their emotions in it. *Phenomenology and the Cognitive Sciences, 11*(2), 135-148.

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

Thompson, E., & Cosmelli, D. (2011). Brain in a vat or body in a world? Brainbound versus enactive views. *Philosophical Topics, 39*(1), 163-180.

Thompson, E., & Varela, F. J. (2001). Radical embodiment: Neural dynamics and consciousness. *Trends in Cognitive Sciences, 5*(10), 418-425.

Varela, F. J. (1999). Present-time consciousness. *Journal of Consciousness Studies, 6*(2-3), 111-140.

Varela, F. J. (2000). Steps to a science of inter-being: Unfolding the Dharma implicit in modern cognitive science. In G. Watson, S. Batchelor & G. Claxton (Eds.), *The Psychology of Awakening: Buddhism, Science, and our Day-to-Day Lives* (pp. 71-89). Boston, MA: Weiser Books.

Wiltshire, T. J., Lobato, E. J. C., McConnell, D. S., & Fiore, S. M. (2015). Prospects for direct social perception: A multi-theoretical integration to further the science of social cognition. *Frontiers in Human Neuroscience, 8*(1007). doi: 10.3389/fnhum.2014.01007

Zahavi, D. (2016). You, me, and we: The sharing of emotional experiences. *Journal of Consciousness Studies, 22*(1-2), 84-101.