

Sociality and the life-mind continuity thesis

Tom Froese and Ezequiel A. Di Paolo

Centre for Computational Neuroscience and Robotics (CCNR)
Centre for Research in Cognitive Science (COGS)
University of Sussex, Brighton, UK

E-mail: t.froese@gmail.com; ezequiel@sussex.ac.uk

Abstract. The life-mind continuity thesis holds that mind is prefigured in life and that mind belongs to life. The biggest challenge faced by proponents of this thesis is to show how an explanatory framework that accounts for basic biological processes can be systematically extended to incorporate the highest reaches of human cognition. We suggest that this apparent ‘cognitive gap’ between minimal and human forms of life appears insurmountable largely because of the methodological individualism that is prevalent in cognitive science. Accordingly, a two-fold strategy is used to show how a consideration of sociality can address both sides of the cognitive gap: (i) it is argued from a systemic perspective that inter-agent interactions can extend the behavioral domain of even the simplest agents, and (ii) it is argued from a phenomenological perspective that the cognitive attitude characteristic of adult human beings is essentially intersubjectively constituted, in particular with respect to the possibility of perceiving objects as detached from our own immediate concerns. These two complementary considerations of the constitutive role of inter-agent interactions for mind and cognition indicate that sociality is an indispensable element of the life-mind continuity thesis and of cognitive science more generally.

Note: This manuscript is the final version of a paper that is going to appear in the journal *Phenomenology and the Cognitive Sciences* (2009).

1. Introduction

The focus of this paper is on a particularly radical element of the recent embodied turn in cognitive science, namely the *life-mind continuity thesis*. This thesis has been proposed in a wide variety of formulations (e.g. Maturana & Varela 1987; Godfrey-Smith 1996; Stewart 1996; Wheeler 1997; Zlatev 2002; Thompson 2007; Di Paolo 2009), but most of them essentially revolve around what has been called ‘strong’ continuity, i.e. that life and mind have a common set of basic organizational properties:

In more concrete terms, the thesis of strong continuity would be true if, for example, the basic concepts needed to understand the organization of life turned out to be self-organization, collective dynamics, circular causal processes, autopoiesis, etc., and if *those very same concepts and constructs* turned out to be central to a proper scientific understanding of mind. (Clark 2001, p. 118)

This strong version is especially attractive for embodied and dynamical approaches to cognitive science for obvious reasons: for if the thesis turns out to be correct then the applicability of these approaches is not only limited to mere low-level, ‘implementation’ details of adaptive behavior. Instead, they would actually be providing the very foundations of a general theory of mind and cognition, one that also includes the highest reaches of human cognition (cf. Clark 2001, pp. 128-130). In other words, the strong continuity thesis aspires to become a unified theory of life and mind.

The most comprehensive theoretical framework based on this thesis is currently being developed by the enactive approach to cognitive science (e.g. Stewart, in press; Di Paolo, et al., in press; Thompson 2007; 2004; Froese 2009). The enactive approach is just as interested in the single-cell organism, i.e. the paradigmatic case of individual agency, as it is in adult human existence, i.e. the paradigmatic case of enculturation. Moreover, as Thompson (2007, p. 129) points out, this approach goes further than systemic life-mind continuity theories by following Jonas’ phenomenological claim that certain basic concepts that are needed to understand human experience turn out to be applicable to life itself as well:

The great contradictions which man discovers in himself – freedom and necessity, autonomy and mortality – have their rudimentary traces in even the most primitive forms of life, each precariously balanced between being and not-being, and each already endowed with an internal horizon of ‘transcendence’. (Jonas 1966, p. ix)

Jonas’ crucial insight is that the phenomenon of life incorporates both living (material) and lived (experiential) aspects. It follows that the enactive version of the strong thesis is not only interested in systemic and behavioral continuity, but also in how this life-mind continuity can be understood in a *phenomenological* manner¹. The study of life can thus

¹ Of course, strictly speaking phenomenology is concerned with one’s own lived experience, so it might appear strange that we want to generalize some of its insights to life as such. It is beyond the scope of this paper to argue for this claim more fully, except to point out that we do have second-person access to the

be seen to comprise not only systemic biology and the philosophy of the organism, but also cognitive science and philosophy of mind.

1.1 Two apparent problems with the life-mind continuity thesis

The development of this kind of radical life-mind continuity thesis as a unified theory of life and mind is not without its problems. Clark, for example, voices some considerable concerns: “The danger, of course, is that by stressing unity and similarity we may lose sight of what is special and distinctive. Mind may indeed participate in many of the dynamic processes of life. But what about our old friends, the fundamentally reason-based transitions and the grasp of absent and the abstract characteristic of advanced cognition?” (Clark 2001, pp. 118-119). We can unpack Clark’s concerns into two related but distinct aspects, namely the problem of agency and the cognitive gap.

On the one hand, there is the morally and scientifically motivated worry that the life-mind continuity thesis “threatens to eliminate the idea of purposive agency unless it is combined with some recognition of the special way goals and knowledge figure in the origination of some of our bodily motions” (Clark 2001, p. 135). In response to this concern it is important to emphasize that the enactive approach is acutely aware of the problem of agency, and most of its efforts are directed toward gaining a better understanding of this phenomenon (cf. Di Paolo 2009; Barandiaran et al, forthcoming; Moreno & Etxeberria 2005). Indeed, the very turn toward the life-mind continuity thesis is largely motivated by a perceived lack of any coherent notion of autonomous agency in current cognitive science. While a full response by the enactive approach to this concern cannot be developed within the scope of this paper (cf. Thompson 2007), its basic theoretical framework of autonomous agency is briefly introduced in Sections 2.1 and 2.2. These sections should alleviate any concerns about whether the idea of purposive agency is in danger of being overlooked.

However, there still remains another, more significant problem that is closely associated with the life-mind continuity thesis: “What, in general, is the relation between the strategies used to solve basic problems of perception and action and those used to solve more abstract or higher level problems?” (Clark 2001, p. 135). Is it a question of merely adding more complexity, i.e. of just having more of the same kind of organizations and mechanisms? Then why is it seemingly impossible to properly address the hallmarks of human cognition with only these basic biological principles? In a recent paper, De Jaegher and Froese (in press) have referred to this missing link as the “cognitive gap” of the life-mind continuity thesis. They propose that this gap is a symptom of the prevalent methodological individualism of cognitive science (cf. Boden 2006), i.e. an exclusive focus on individual agency whereby all social phenomena, including social cognition, is reducible without remainder to individual mechanisms, and that the gap can therefore be addressed by taking the role of sociality into account.

experience of others (cf. Section 3), and that this can evidently include some other forms of life. Moreover, we can correlate our experience with our biological conditions and this enables us to justify from the third-person perspective that other organisms are likely to undergo similar experiences.

A similar response has been developed by ‘extended mind’ theorists such as Clark, who proposes “to depict much of advanced cognition as rooted in the operation of the same basic kinds of capacity used for on-line, adaptive response, but tuned and applied to the special domain of *external and/or artificial cognitive aids*” (2001, p. 141). However, these efforts have largely focused on the role of language (e.g. Clark 2008, pp. 44-60) and technology (e.g. Clark 2003), thereby relating specifically human cognition with specifically human abilities and their cultural context (for related proposals, cf. Stewart (in press) and Zlatev (2002)). Thus, while this consideration of ‘cognitive technology’ helps to spread the explanatory burden outside of the individual human agent, and thereby makes the basic principles of embodied-embedded accounts more plausible, it still leaves the cognitive gap of the life-mind continuity thesis largely unaddressed.

1.2 The constitutive role of sociality

An extended or distributed view of cognition complements the life-mind continuity thesis nicely because it can show how complex behavior does not necessarily have to be the result of similarly complex internal processes. As such, giving up the traditional internalist doctrine is indeed a crucial first step to make the strong thesis plausible. But this extended view is also in need of further explication as it entails nothing more than a general commitment to the thesis that cognition emerges out of the dynamics of a brain-body-world systemic whole (e.g. Beer 2000). What is additionally needed is a more detailed non-species-specific operational mechanism to account for the transformative potential of such cognitive extension as it relates to agency (Di Paolo 2009). To be sure, the desirability of a more encompassing theoretical account is not denied by extended mind proponents. Clark (2005), for example, suggests that the sound-amplifying burrow of the mole cricket is a loose analogy to the cognition-transforming symbols found in human culture. But it is important to note that the chirping cricket in its burrow is *passively* interacting with a *static* physical structure. Moreover, this example completely ignores the fact that human symbols only exist within a *social* context.

Accordingly, De Jaegher and Froese (in press) would agree with Clark that “interactive complexity characterizes almost all forms of advanced human cognitive endeavor” (Clark 2001, p. 154), but they argue that such interactive complexity is already prefigured in the social co-constitution of more basic cognitive domains. It has been demonstrated in the field of artificial life that even interactions between minimal kinds of ‘agents’ can give rise to interaction processes that are characterized by autonomous dynamics, i.e. which self-sustain their identity as processes by modulating the behavior of the interactors in an appropriate manner (e.g. Froese & Di Paolo, in press; 2008). Here we have an example of cognitive extension that involves *active* coordination, *dynamic* emerging structures, and an interactive *social* context that removes the need for static physical structures. In the enactive approach to social cognition this transformative potential of the autonomous interaction process has been termed “participatory sense-making” (De Jaegher & Di Paolo 2007). The basic ideas behind this approach are introduced in Section 2.3.

However, by concentrating on the role of interaction dynamics in this *systemic* manner, the *meaning* of the presence of another agent for an individual’s perspective on the world

has largely remained neglected so far. In other words, De Jaegher and Froese (in press) focus on the organizational and behavioral side of sociality, but not so much on the related phenomenological implications. Accordingly, one of the aims of this paper is to complement their development of the life-mind continuity thesis through an analysis of the phenomenology of intersubjectivity (Section 3). It is argued that our human first-person perspective, with its distinct capacity for detached reflection and abstract reasoning, is fundamentally intersubjectively constituted. This transformative potential of intersubjectivity gives additional support to the strong continuity thesis: for it suggests that it is not any dedicated internal mechanism that makes these ‘advanced’ forms of cognition possible, but rather our particular way of interacting with others. The main conclusion of this phenomenological analysis therefore complements the work by De Jaegher and Froese (in press) by indicating that the life-mind continuity thesis can indeed become a viable contender for a unified theory of life and mind – but only when we take the constitutive role of sociality into account.

2. The enactive approach to agency and sense-making

Before we consider how sociality impacts on enactive cognitive science’s conception of autonomous agency and sense-making, it is first necessary to properly define these two notions. However, since these notions have already been the target of much recent work (e.g. Di Paolo 2009; Thompson 2007; Froese & Ziemke 2009), here we only highlight some aspects. In particular, the focus is on the mechanisms of identity generation and adaptive regulation which, in combination, result in the enaction of a basic perspective for which events show up as significant, i.e. as meaningful in relation to the norms that are established by the self-conservation of the underlying autonomous identity. The aim of this section is to show that the life-mind continuity thesis does not undermine the notion of agency, but rather grows out of a systematic consideration of the notion’s biological foundation.

2.1 Autopoiesis – intrinsic teleology

The development of the theoretical framework of the enactive approach has its historical beginnings in the early 1970s, in the work of the Chilean biologists Maturana and Varela (e.g. Varela, et al. 1974; Maturana & Varela 1987). Dissatisfied with mainstream biology, they instead followed in the footsteps of the cybernetics movement and came up with a systemic definition of the autonomous organization of the living, which they termed *autopoiesis*. The precise definition of this term has changed slightly over the years, and the inherent circularity of the minimal living organization makes its proper definition rather lengthy, but basically the notion of autopoiesis contains two essentially interrelated aspects: systemic identity-generation and material self-production. In other words, the autopoietic system is organized such that its ongoing operations have the effect that it continually produces itself materially. Moreover, this self-production entails that the system can autonomously generate its own identity. It is important to realize that, by generating its own identity in this manner, the autopoietic system simultaneously generates the particular conditions by which it can relate to its environment. Indeed, it is

this fundamental *asymmetry* of the organism-environment relationship, which partly constitutes the organism's *perspective* on that environment.

Of course, this is not to say that we cannot provide a description of the organism and its environment in physicochemical terms. The point is merely that this type of description does not exhaust the domain of phenomena with which biology should be concerned. In other words, "one could envisage the circularity metabolism-membrane entirely from the outside (this is what most biochemists do). But this is not to deny that there is, at the same time, the instauration of a *point of view* provided by the self-construction" (Weber & Varela 2002, p. 116). More precisely, this self-generation of a perspective by which events show up as meaningful for the organism consists in two essential and inseparable aspects: (i) the autonomous generation of an identity, and (ii) the constitution of a relationship to what is other, whereby the norms established by the self-conservation of the former act as a reference point for the latter (Weber & Varela 2002).

Let us consider these two aspects in turn. First, there is the notion of *intrinsic teleology* in terms of the organism's relation to its own identity, i.e. its capacity to constitute its own purposeful and goal-directed existence. Weber and Varela begin to derive this notion by combining Kant's conception of a *natural purpose*, namely the idea that a self-organizing system that is both cause and effect of itself is also its own means and purpose (cf. Kant 1790, §64-65), with our modern understanding of autopoiesis. Then, by appealing to the philosophical biology of Jonas, they move beyond Kant's conception of teleology as a useful regulative idea for the observer, and posit teleology as *intrinsic* to the phenomenon of life itself. Indeed, Jonas himself argues that the precarious situation of the living furnishes the organism with more than just goal-directed behavior (a type of behavior which can be ascribed to any system characterized by negative feedback); it also entails a fundamental *existential* concern: "The organism has to keep going, because to be going is its very existence – which is revocable – and, threatened with extinction, it is *concerned* in existing" (Jonas 1966, p. 126, emphasis added). It is this existential concern for the self-preservation of its precarious identity that gives the organism an intrinsically purposeful and meaningful perspective on encountered events. Jonas thus proposes that it is the generation of a precarious identity through self-production that simultaneously enables the constitution of *values* for that identity:

The basic clue is that life says yes to itself. By clinging to itself it declares that it values itself. [...] Are we then, perhaps, allowed to say that mortality is the narrow gate through which alone *value* – the addressee of a yes – could enter the otherwise indifferent universe? (Jonas 1992, p. 36)

Jonas' consideration of the metabolic origin of values brings us to the second aspect of the organism's perspective, namely its capacity for sense-creation or *sense-making*. This notion highlights that the constitution of values always happens in the context of a particular asymmetric organism-environment relationship that is centered on the identity of the organism. The internal and external encounters that perturb the precarious process of identity generation thereby take on a significance in relation to the effects of this perturbation: "The perspective of a challenged and self-affirming organism lays a new

grid over the world: a ubiquitous scale of value.” (Weber & Varela 2002, p. 118). This notion of enaction as sense-making can be nicely illustrated by one of Varela’s favorite examples: “There is no food significance in sucrose except when a bacteria swims upgradient and its metabolism uses the molecule in a way that allows its identity to continue” (Varela 1997, p. 79). In other words, an organism’s world is primarily a context of significance in relation to that organism’s particular manner of realizing and preserving its precarious identity.

2.2 Adaptivity – sense-making

Let us take a closer look at the second aspect of an organism’s perspective, namely its sense-making capacity. Varela has argued that “the source for this world-making is always the breakdowns in autopoiesis” (Varela 1997, p. 80). However, as Di Paolo (2005) has recently pointed out, the original concept of autopoiesis as developed by Maturana and Varela allows no gradation by itself – either a system belongs to the class of such systems or it does not (cf. Varela 1997, p. 77). The self-constitution of an identity can thus provide us only with the most basic kind of norm, namely that all events are good for that identity as long as they do not destroy it (and the latter kind of events are not significant because there will be no more identity to which they could be related). On the basis of autonomous identity-generation alone it is impossible to account for the different shades of meaning which are constitutive of an organism’s world. Even the humble sucrose-seeking bacterium must be able to distinguish sucrose gradients from other non-significant factors. But how can we extend the meaningful perspective that is engendered by constitutive autonomy into a wider context of relevance?

Di Paolo (2005) has proposed a resolution to this problem which builds on the all-or-nothing concept of autopoiesis but nevertheless allows for the constitution of different kinds of meaning. As a first step he defines an autopoietic system’s range of non-fatal events as its viability set. In order for an autopoietic system to actively improve its current perturbed and precarious situation, it must (i) be capable of determining how the ongoing structural changes are shaping its trajectory within this viability set, and (ii) have the capacity to regulate the conditions of this trajectory appropriately. Adaptivity, as defined in this manner, is crucial for our understanding of an organism’s perspective for two reasons: (i) it allows the system to distinguish between some positive and negative tendencies, and (ii) it ensures that the system can measure the type and severity of such a tendency according to a change in the regulative resources required.

It is important to note that the capacity for (i) does not contradict the autonomy of the autopoietic system because it depends on (ii). In other words, the system does not have any special epistemic access to an independent (non-relational) environment, and it therefore does not violate the relational nature of constitutive autonomy; this is not a problem since the system only needs to monitor *internal* regulatory effort. Furthermore, it is worth emphasizing that (ii) already implies the capacity for suitable compensation. If there are perturbations for which no adaptive response is available (e.g. radioactive radiation) they do not show up as significant for the organism. Thus, “if autopoiesis in the present analysis suffices for generating a natural purpose, adaptivity reflects the

organism's capability – necessary for sense-making – of evaluating the needs and expanding the means towards that purpose" (Di Paolo 2005, p. 445). We can therefore say that adaptivity, i.e. self-monitoring and appropriate regulation, is necessary for the enaction of a perspective for which different kinds of meaning can show up.

We can summarize these biological foundations of enactive cognitive science in the form of two systemic requirements. First, it is claimed that the self-generating identity of autopoiesis (or constitutive autonomy more generally, cf. Froese, et al. 2007) is necessary for intrinsic teleology and sense-making. Second, it is claimed that the property of adaptivity is also necessary for sense-making. Finally, for the purposes of this paper, we define any system which meets these two requirements, and for which its sense-making extends to the active domain of interactions, as an autonomous agent (see also, Barandiaran, et al., forthcoming).

2.3 Social interaction – participatory sense-making

It is evident that our lived perspective, and presumably that of most other living beings as well, is characterized by a whole range of different shades of meaning. This meaningful differentiation has to be matched in operational terms by a notion capable of giving rise to such graded distinctions. The concept of adaptivity is a first step in this direction because it has enabled us to become clearer about what we mean by the notion of sense-making. However, this conceptual advance is just the beginning of a new dialectic between the identification of experiential phenomena and the formulation of operational mechanisms that can account for them. Indeed, the term's general applicability to all essentially all living beings cries out for further specification that can distinguish between different modes of sense-making.

In this paper, however, we are specifically concerned with De Jaegher and Froese's (in press) proposal that one way to begin to account for the cognitive gap of the life-mind continuity thesis, i.e. the difference between the sense-making capabilities of a simple single-cell organism and that of a fully developed human agent, is by acknowledging the constitutive role of inter-individual interaction. How should such an interaction be characterized? Building on the notion of constitutive autonomy as the self-generation of an identity under precarious conditions, it is possible to propose the following definition:

Social interaction is the regulated coupling between at least two autonomous agents, where the regulation is aimed at aspects of the coupling itself so that it constitutes an emergent autonomous organization in the domain of relational dynamics, without destroying in the process the autonomy of the agents involved (though the latter's scope can be augmented or reduced). (De Jaegher & Di Paolo 2007, p. 493)

There has been some debate within enactive cognitive science whether this definition manages to capture what is specific about sociality (e.g. Steiner & Stewart, in press). But even if it turns out that the kind of multi-agent interaction that De Jaegher and Di Paolo describe is not a *sufficient* condition for *social* interaction, this would only strengthen the

bottom-up continuity argument of this paper. In fact, just as the notion of autopoiesis was complemented by adaptivity to account for autonomous agency, we can expect that future work in this area will further explicate some additional necessary conditions for social interaction. These could then help us to specify more precisely the various constitutive contributions of inter-individual interactions for individual agency and sense-making.

For the purposes of this paper, however, De Jaegher and Di Paolo's definition is a good starting point. Indeed, with respect to the goal of further developing the notion of sense-making so as to address the cognitive gap of the life-mind continuity thesis, two aspects of their definition are particularly noteworthy: (i) since the interacting agents are autonomous systems (constitutive autonomy) and they regulate this interaction (adaptivity), it follows that they can engage with each other in terms of sense-making, and (ii) since the regulation of the interaction by one agent changes not only its own coupling but also that of the other agent, it follows that the agents can enable and constrain each other's sense-making:

If regulation of social coupling takes place through coordination of movements, and if movements – including utterances – are the tools of sense-making, then our proposal is: social agents can coordinate their sense-making in social encounters. [...] This is what we call *participatory sense-making*: the coordination of intentional activity in interaction, whereby individual sense-making processes are affected and new domains of social sense-making can be generated that were not available to each individual on her own. (De Jaegher & Di Paolo 2007, p. 497)

It is worth emphasizing the basic idea of this proposal again: if agents mutually enable and constrain their sense-making activities in an appropriate manner, they can open up new domains of sense-making that would have otherwise remained inaccessible to the individual agents. Accordingly, De Jaegher and Froese (in press) argue that the notion of participatory sense-making can help us to systematically address the cognitive gap from the bottom-up. In support of this view they cite some recent experimental and modeling work in the cognitive sciences.

An important example is the recent psychological study by Auvray and colleagues (2009) who investigated the role of inter-individual interactions in organizing appropriate task-solving behavior under minimal experimental conditions. Two participants were asked to locate each other in a 1-D virtual environment. The participants are able to control their avatar's movement in the virtual space; they receive an all-or-nothing tactile feedback to the finger whenever their avatar is overlapping any other object within that space. They can encounter three types of objects: (i) a static object, (ii) the avatar of the other participant, and (iii) a 'shadow' copy of the other's avatar that exactly mirrors the other's movement at a displaced location. The task is non-trivial because all of these objects are of the same size and thus generate the same all-or-nothing tactile response. The only way to differentiate between the objects is through the distinct kinds of interaction dynamics that can result from their encounter. Nevertheless, participants did manage to locate each other successfully most of the time, despite the minimal conditions under which they had to complete the task. This is because ongoing perceptual crossing, i.e. when both avatars

mutually interact with each other, affords the most stable interactive situation under these circumstances. The surprising result of the study is that even though the participants ‘failed’ to achieve the task individually, i.e. there was no significant difference between their probabilities of recognizing the other’s avatar and the other’s shadow object, they managed to solve the task collectively, finding most of the time the other’s avatar and ignoring their shadow (cf. Auvray, et al. 2009, p. 39).

Similarly, in the field of artificial life it has been shown that it is possible for a dynamical system to distinguish between interactions with another ‘live’ partner or a non-responsive recording, i.e. what in psychology is known as sensitivity to social contingency, by means of the interaction process itself (e.g. Di Paolo, et al. 2008). It has also been demonstrated that two dynamical systems can mutually enable and constrain each other’s behavior in such a way that this leads to new behavioral flexibility, for example increased capabilities for movement (e.g. Froese & Di Paolo 2008; in press). These rather brief descriptions cannot do justice to the detailed understanding of the dynamics of inter-agent interactions, including their constitutive role for increased individual behavioral domains, which these and other models can provide. However, unfortunately it is beyond the scope of this paper to address these modeling implementations and results more specifically; the reader is therefore encouraged to follow up the references provided here, as well as the more comprehensive analysis provided by De Jaegher and Froese (in press).

In sum, these kinds of experiments provide important support for the proposal that inter-individual interaction can extend the *behavioral* repertoire of the interacting agents. It is thus indeed possible to begin addressing the cognitive gap from the bottom-up in this manner. However, does the behavioral extension afforded by this interaction provide us with a specific *qualitative* change in sense-making? That is, does the observation of such interaction dynamics between two autonomous agents allow us to conclude that their *world* is qualitatively different from that of an individual agent?

2.4 Accounting for the presence of the other

At first sight this might appear to be obviously the case – after all, are we not precisely talking about participatory sense-making as opening up new *social* domains of sense-making (cf. De Jaegher & Di Paolo 2007, p. 497)? However, what these experiments and models have also shown is that De Jaegher and Di Paolo’s definition of social interaction can apply to the regulated interaction between two autonomous agents regardless whether or not the individual agents are even *aware* of the social nature of this interaction. In other words, the interaction process between two agents can enable and constrain an individual agent’s sense-making capabilities without the other agent being necessarily present within that individual’s world *as another agent*. On this view, it is possible that an agent enacts a qualitatively different world when it happens to do so within a multi-agent system, but as long as other agents are not significant as such these qualitative changes do not include those associated with interactions in a social domain.

For example, as the artificial life models have demonstrated, autonomous interaction dynamics can already occur between extremely simple dynamical systems that are

themselves neither constitutively autonomous nor adaptive, and which therefore have no sense-making capacities at all. To be sure, the point of these models was precisely to provide *models* of the interaction process, rather than to replicate its natural conditions of emergence, but the possibility of doing so under such minimal and abstract conditions does say something about the potential generality of this phenomenon. Moreover, while the interaction process in the experimental setup of Auvray and colleagues enables the individual participants to successfully achieve the task together, this joint success is only verifiable from the privileged perspective of an outside observer. Of course, all of this is not to say that the kind of interaction dynamics that have been specified by De Jaegher and Di Paolo do not also occur during social encounters where the other agent is actually experienced as another agent; they most likely do. Nevertheless, it is important to realize that the current operational definition of “social interaction” does not enable us to distinguish between these different kinds of situation. In other words, the relationship between the operational definition and the phenomenological concept of intersubjectivity remains unclear. So what is missing?

At this point we will make use of what has almost become a trademark move of the enactive approach to cognitive science. Recall that it was our own lived experience of leading purposeful and concerned existences that made Weber and Varela, following Jonas, extend the concept of autopoiesis in its current direction. Similarly, it is our experience of a world that is characterized by different shades of meanings that motivated Di Paolo in proposing the principle of adaptivity as a mechanism which could realize such differentiation. And now we have come across another particular experiential phenomenon, namely the encounter with another subject in its own right, which is unaccounted for in operational terms. Accordingly, in order to better understand precisely what this phenomenon of the presence of the other amounts to, we will consider some of the literature on the phenomenology of intersubjectivity (Section 3). This will not only provide additional support for the constitute role of sociality and the life-mind continuity thesis, but also the basis for future attempts at relating systemic and phenomenological accounts of the social in a mutually informing and constraining manner. Some initial steps in that direction will be presented at the end of this paper (Section 4).

3. Phenomenology of intersubjectivity

In this section we will draw on some central insights from Husserlian phenomenology, especially in relation to the problem of intersubjectivity which has taken center stage for much of the tradition’s later developments. For an extensive discussion of the relevant phenomenological literature the reader is referred to Zahavi’s (1996) excellent treatment of the topic, to which this section owes much. While there is a growing recognition of the importance of a phenomenologically informed approach to intersubjectivity for the cognitive sciences (e.g. Gallagher & Zahavi 2008; Gallagher 2008a; Zahavi 2005; 2001), and the enactive approach in particular (e.g. Thompson 2001; 2007), these are only the tentative beginnings of a mutually informative research program. In the context of this paper, we will limit our analysis to a consideration of how the phenomenology of intersubjectivity can contribute to a better understanding of the life-mind continuity

thesis. Of particular interest will therefore be the development of insights into how the presence of others impacts on the structures of agency and sense-making.

This section will unfold in two parts. First, we will review aspects of the phenomenology of intersubjectivity in relation to perception, which is how Husserl first got drawn into a serious appreciation of the constitutive role of intersubjectivity, and which has also recently been used to motivate a similar turn in the cognitive sciences (e.g. Gallagher 2008a; Zahavi 2005, pp. 166-167). Second, we will clarify the way in which others are given to us in our experience, giving special attention to how this encounter affects our own perspective on the world. These phenomenological reflections indicate how the objectivist epistemic perspective that is characteristic of detached human sense-making (e.g. the scientific attitude) is constitutively dependent on our de-centering relationship with others. In Section 4 the life-mind continuity thesis is reformulated to include this phenomenological support for the constitutive role of sociality.

3.1 Intersubjectivity in the phenomenology of perception

One of the key insights of the phenomenology of intersubjectivity is that the existence of others plays a constitutive role for our perception. In order to illustrate this insight, let us begin with a concrete example of object perception: how is the wall that is located behind the desk given to me in my experience? How is its presence for me as an object constituted? The general proposal of the enactive approach to perception is that the world (our experiential world) is *enacted*; that is, it claims that perception consists in perceptually guided action. The perceiver does not have access to some pre-given, perceiver independent world of objects since an autonomous agent is always embedded within a particular sensory-motor loop that is shaped by the overall dynamics of the organism-environment systemic whole (Varela, et al. 1991, p. 173). Accordingly, such an agent can only perceive its surroundings by regulating its sensory-motor interactions.

On this view, perceived objects appear as the invariants that emerge from the closed loop of an agent's ongoing embodied activity and the resulting stimulations (an idea inherited from the cybernetic tradition, e.g. von Foerster 1976). It is important to emphasize that we are specifically talking about *sensory-motor invariants*, which differ from the ecological invariants of Gibsonian psychology in that only the former give a constitutive role to motor activity (cf. Mossio & Taraborelli 2008). Since it is my behavior which allows me to establish such sensory-motor regularities, it follows that my perceptual capabilities are enabled and constrained by my behavioral capabilities:

How things look *to me* is constrained by my sensorimotor knowledge. It is my possession of basic sensorimotor skills (which include the abilities to move and point and the dispositions to respond by turning and ducking, and the like) that enables my experience to acquire *visual* content at all. (Noë 2004, p. 90).

Thus, according to the enactive approach, my perception of the wall behind my desk is constituted by my possession of basic sensory-motor skills, such as perhaps scanning it with my eyes or moving my head and body in a certain manner. This sensory-motor

account of perception has generated a lot of excitement in cognitive science, but also a number of criticisms. These need to be taken seriously and carefully addressed in order to move the theory of sense-making and enactive perception forward (e.g. Di Paolo 2009). Here we will focus on one problematic aspect for postulating a sensory-motor basis for perception, namely the perceptual presence of absent phenomena, or what is sometimes called the problem of “perceptual presence” (Noë 2004, p. 59). This problem is of particular interest in the current context because a related worry in phenomenology has been resolved by appeal to the constitutive role of open intersubjectivity.

Let us return to the example of the wall behind the desk. How come it is given to me in my perceptual experience as a 3D object that has another side which is currently out of view? How come it is given to me as separating my current location from whatever is outside the room, rather than just as a flat 2D appearance? Noë suggests that our experience of absent profiles should be understood as a type of ‘virtual’ presence: “They are present to perception *as accessible*” (Noë 2004, p. 63; emphasis added). In other words, on this account I do not experience the wall as a flat facade that separates me from some meaningless void because of my embodied sensory-motor skills, which would let me view its other side from the outside, if I left the room to inspect it.

But is this appeal to ‘virtual’ presence in terms of potential sensory-motor accessibility an adequate description of our perceptual experience? According to the later work of Husserl and the phenomenological tradition, as well as recent work in cognitive science, this is not the case (cf. Gallagher 2008a; Zahavi 1996, pp. 43-51; Gallagher & Zahavi 2008, p. 100-104; Thompson 2007, p. 384). The idea that we constitute the absent profiles of an object in terms of past or future embodied action necessitates an appeal to a *temporal* separation, an aspect that is itself not given in our perceptual experience. Even though I momentarily do not visually perceive the backside directly, I nevertheless experience the wall as having such a backside *now* while I am looking at it from inside my room. In other words, it is argued that since leaving the room to check the wall’s backside would involve a temporal and spatial displacement from my current perceptual situation, neither the movement nor its potential accessibility can explain the fact that I *currently* perceive the wall as a whole object, no matter which side aspect is given to me, rather than as a temporarily distributed set of profiles. As Gallagher, following Husserl, points out: “When I perceive an object the present front is not a front with respect to a past or future back, but is determined through its reference to a present co-existing back. The object is perceived at any given moment as possessing a plurality of *co-existing* profiles” (Gallagher 2008a, p. 172). But since I can only be in one place at a time, how then can we account for this perceptual presence of co-existing profiles?

We can begin to resolve this problem for the enactive approach by noting that Varela and colleagues make use of a much broader notion of embodiment than that used by Noë. The term ‘embodied action’ is indeed meant to highlight that cognition and perception depend on having a biological body which entails various kinds of constraints and skills. But the term is similarly meant to emphasize “that these individual sensorimotor capacities are themselves embedded in a more encompassing biological, psychological, and cultural context” (Varela, et al. 1991, pp. 172-173). To be sure, so far much of enactive cognitive

science has focused on the biological and psychological context, but the tradition has also always been informed by considerations of our social and cultural background, especially in terms of our role as practicing scientists (cf. Varela, et al. 1991, pp. 9-12; Thompson 2007; 2001). Moreover, even in the very beginning of the enactive approach the role of social context enters into the very definition of what it means to be an intelligent agent, such that “intelligence shifts from being the capacity to solve a problem to the capacity to enter into a *shared* world of significance” (Varela, et al., p 207; emphasis added). How does this broader notion of embodiment, which includes our situatedness in a social and cultural context, help us to account for our perception of the wall as a full-fledged object?

First, following Gallagher (2008a, p. 171), we can note that there is good evidence from developmental psychology that we gain access to a meaningful world of objects through our interactions with others. Not only are the most dominant and central experiences for a young infant its relations to others, these relations also shape its perception of the world by engaging in joint attention. In other words, environmental objects, such as a wall, first take on meaning in the pragmatic contexts within which we see and imitate the actions of others (cf. Merleau-Ponty 1960; Trevarthen & Hubley 1978; Tomasello 1999). Second, even in our adult life we find that the wall is given as ‘a wall’ within a common *public* totality of surroundings, i.e. in phenomenological terms the situation of our individual being is always already a form of ‘being-with others’ (cf. Heidegger 1927; Zahavi 1996, pp. 124-127). Similarly, research in cognitive anthropology has demonstrated that our relations to the objects of our perception continue to be deeply interwoven with our inter-individual relations to others even in adult life (cf. Hutchins 1995). Thus, the capacity for worldly engagement that is characteristic of adult humans is neither acquired nor performed in isolation. As Husserl famously puts it:

Thus everything objective that stands before me in experience and primarily in perception has an apperceptive horizon of possible experience, own and foreign. Ontologically speaking, every appearance that I have is from the very beginning a part of an open endless, but explicitly realized totality of possible appearances of the same, and the subjectivity belonging to this appearance is open intersubjectivity. (Hua XIV/289; see also Hua IX/294, XV/497; quoted by Zahavi 2005, p. 167)

In this manner Husserl was led via a deepening phenomenological analysis of object perception from an individualistic theory of the constitution of objects, which is akin to the modern sensory-motor approaches to perception, to an appreciation of the constitutive role of other subjects. More precisely, he began to realize that in order to account for the phenomenology of an object as something that transcends the current perceptual profile that we have of it, we need to posit the possibility of other subjects, which can potentially perceive the other profiles at the same time.

But this intersubjective resolution of the problem of ‘perceptual presence’ leaves us with a dilemma with regard to the life-mind continuity thesis. For if we accept that the full presence of the wall as a 3D object is constituted via a specific form of intersubjectivity, namely in terms of the potentiality of my experience of others *as* others that can perceive

its currently hidden profiles, then what does this entail for other forms of life? We seem to have two controversial options: (i) we claim that such open intersubjectivity is present for other living beings as well, or (ii) we assert that the presence of a 3D world is a unique aspect of human phenomenology. While the former claim remains highly contentious even with regard to our closest primate relatives (Tomasello, et al. 2003) and lacks scientific support for most other species (Tomasello 1999), the latter is in tension with the evident skill with which these species negotiate spatial environments. Should we really say of those species that do not understand others *as* others like themselves that they are lacking the experience of whole perceptual objects in a spatial world? That seems to be a rather absurd consequence. Thus, even if it turns out that the required social understanding is more widespread than is currently assumed, it appears that a more parsimonious explanation for such a basic experiential category is desirable. We propose to resolve this dilemma in two steps.

First, let us reconsider the constitution of perceptual presence in terms of sensory-motor invariants. In humans, at least, minimalist psychological experiments have shown that the emergence of an experience of distal presence (in terms of an obstacle located in space) only requires the mastery of some basic sensory-motor correlations and laws (Auvray, et al. 2005). Since we are here dealing with the constitution of a completely novel action-space simply by means of an individual's (technologically mediated) exploration, the role of open intersubjectivity for object perception has been further marginalized. However, we also know that sensory-motor invariants include a temporal progression which makes it impossible to account for the perceptual presence of a complete object by means of its current profile, because "the object is perceived at any given moment as possessing a plurality of *co-existing* profiles" (Gallagher 2008a, p. 172). We suggest a novel resolution of the tension between these two accounts by appealing to another foundational theme of phenomenology and enactive cognitive science, namely the issue of temporality (cf. Varela 1999). Husserl himself has argued that every experiential moment is not simply an isolated point in time, but is rather temporally extended in terms of a tripartite retention-present-protection structure (cf. Hua X). In view of this, it must be clarified that the sensory-motor view does not require the co-existence of any two actualities, which would indeed demand a separation in time, but the co-existence of an actuality and a virtuality, the possibilities that are open now to me should I decide to take this or that course of action. This virtuality is accessible because our experience of time is extended to an immediate past and projected forward. It is in this protection that the intentionality of my perception extends to non-actualized possibilities, as is the case of intentionality in general, and these give form to my perceptual experience. Accordingly, the temporal horizon of the present moment allows us to account for the constitution of an object as *presently* consisting of a plurality of possible (actual and virtual) perspectives, since these can be presently contained in the present moment's retention and protection, even without the need for open intersubjectivity.

It appears that our resolution has undermined the phenomenology of intersubjectivity in preference for an approach that gives support to methodological individualism. To be sure, with the appeal to the temporal structure of experience we have indicated how it is possible to retain the insights of the sensory-motor approach to perceptual presence

without requiring that other subjects must be potentially be present *as* others. But where does this leave the constitutive role of open intersubjectivity? This is where the second step comes into play. We claim that the aspect of perception which is intersubjectively co-constituted is a certain kind of de-centering or detachment that enables us to specifically perceive something *as* something. What is intersubjectively constituted, in this analysis, is not my perception of an object as having hidden profiles that I apperceive – for indeed, this is what we would expect an animal with a visual system close to ours to also perceive. What is intersubjectively constituted is a “change of meaning” of my perception. For example, I don’t just see the wall from the perspective of my own current concerns (e.g. as trapping me inside the room) but I can see it also from the perspective of alternative concerns (e.g. as blocking the view, making parking difficult, etc) including those others might have. These alternative perspectives on the situation need not be actually thematized; their potentiality as legitimate concerns makes me see the wall *as* ‘a wall’ that is independent from my current perspective of it (construed largely as what the wall means to me now), i.e. as an *object* in the strict sense of the word. Here we need to appeal to the possible presence of others *as* others in order to account for the existence of these potential co-existing perspectives, as well as the corresponding relativization or de-centering of our own current perceptual perspective.

On this phenomenological view, it is open intersubjectivity that co-constitutes the characteristically human de-centered presence, a presence for which the sensory-motor enaction of spatiality is a necessary but not sufficient condition. Unlike other animals, humans are not always captive of the immediate sense of their perception. This appeal to the constitutive role of others, as a means of going beyond the limitations of individual sensory-motor knowledge toward the possibility of adopting different perspectives and alternative meanings to a situation, has occasionally been recognized in sensory-motor approaches to developmental psychology (e.g. Piaget 1967; Donaldson 1992), but has so far been lacking in most enactive cognitive science (and, of course, in mainstream cognitive science more generally). We will return to some of the effects of this ‘de-centering’ presence of the other in the next section, in the context of a more detailed phenomenological investigation of how we perceive others.

3.2 The phenomenology of encountering another subject

How precisely do we encounter another subject *as* another subject? In cognitive science this ‘problem of other minds’ is usually addressed according to the computationalist sense-model-plan-act schema: (i) we first perceive a set of physical facts, such as material bodies and movements; (ii) these facts are the input for some kind of cognitive processing, such as inference or simulation, thereby producing a representation of the other and his intentional states, (iii) that representation allows us to plan how to act in a socially appropriate manner, and (iv) we respond by executing that plan with motor commands as best as possible (for a critical analysis of these kinds of inference and simulation based approaches, cf. Gallagher 2001; Zahavi 2001). Though there are some essential differences between the competing theories of social cognition, they are in agreement that the meaningful expressions of the other are essentially a secondary attribution to the primary perception of merely physical circumstances.

What is most peculiar about this way of approaching the phenomenon of sociality is that it does not match what is actually given in our experience: we directly perceive the other as another subject in its own right without having to engage in inference or simulation in order to perceive her as such (cf. Gallagher 2008b; Zahavi 2001). Of course, sometimes it helps to reason about the actions of others, or to imaginatively ‘put ourselves in their shoes’ so as to better clarify their intentions. But this typically happens only in some exceptional circumstances when normal social interaction breaks down, and even then this does not mean that we question the other’s status as a subject as such. Moreover, if we accept that part of what it means to experience an objective world is that we encounter it as a shared existence for other subjects, i.e. the world is experienced as our common world (Hua I/123; cf. Zahavi 1996, pp. 25-26), then the traditional approach, based as it is on an objectivist premise, presupposes what it sets out to explain. The attempt to reduce the presence of others to a combination of physical facts is doomed to failure because the condition of possibility for those facts includes the presence of others. In brief, it is impossible to conceive of objectivity without positing intersubjectivity at the same time.

However, the phenomenological observation that other subjects are immediately given in our experience should not be misunderstood as asserting that we have direct access to their experience in the same way that they themselves have immediate access. On the contrary, a crucial element that defines another subject as such is the peculiar ‘otherness’ of their presence. Indeed, if this were not so then the phenomenon of intersubjectivity would be logically inexplicable. As Husserl puts it: “if what belongs to the other’s own essence were directly accessible, it would be merely a moment of my own essence, and ultimately he himself and I myself would be the same” (Hua I/139). It is evident that oneself and the other have different modes of access to the other’s unfolding experience, what has sometimes been called second- and first-person access, respectively. To be sure, to simply say that the presence of the other is defined by his otherness is not enough. All real objects of our experience are characterized by a certain type of transcendence in relation to the constituting subject. Moreover, we can find structures of otherness (alterity) even within the subject’s relation to itself. Thus, what is special about the transcending presence of another subject is not that the other, as a real phenomenon, necessarily eludes our grasp, but that it does so in a manner that is unique to encounters with others. Let us unpack this specific otherness of the other in more detail.

Objects elude our grasp, i.e. they are necessarily only given in profiles that never exhaust the constituted object as such. This type of object transcendence can be captured by the notion that there is always more than meets the eye. Nevertheless, we can still posit their identity as an ideal limit point, which we pre-reflectively understand through our mastery of sensory-motor engagement with them². Another subject, in contrast, is always prone to change its identity over time in such a way that it escapes any attempt at grasping it in the

² “The ipseity is, of course, never *reached*: each aspect of the thing which falls to our perception is still only an invitation to perceive beyond it, still only a momentary halt in the perceptual process. [...] What makes the ‘reality’ of the thing is therefore precisely what snatches it from our grasp. The aseity of the thing, its unchallengeable presence and the perpetual absence into which it withdraws, are two inseparable aspects of transcendence” (Merleau-Ponty 1945, p. 271).

form of a simple object perception. As long as the other remains an autonomous subject in its own right, there is always the possibility that the affordances of interaction change in surprising and unexpected ways. In relation to the transcendence of things we can say that “the real lends itself to unending exploration; it is inexhaustible” (Merleau-Ponty 1945, p. 378). To be sure, this is also the case for our encounters with other subjects, but there it turns into what we could call a meta- or *second-order transcendence*: others do not only lend themselves to unending exploration like objects, they also spontaneously lend themselves to unending explorations of *different styles* of unending exploration.

This insistence on the radical otherness of the other might at first seem like a minor technical point, but it actually is at the heart of why a consideration of intersubjectivity is so important for any adequate account of human cognition. It is only in the case when the subject encounters the particular second-order transcendence of the other that we can say that its immanent sphere of ‘ownness’ is surpassed toward a *shared* world of objects (Hua XIV/442). In this way we have turned the traditional problem of other minds on its head: “the otherness of ‘someone else’ becomes extended to the whole world, as its ‘Objectivity’, giving it this sense in the first place” (Hua I/173). We thus find that the categories of transcendence, objectivity and reality are intersubjectively constituted, that is, they can only be constituted by a subject who has experienced and interacted with other subjects. More precisely, these categories are co-constituted, as illustrated by the following example of their application to our understanding of ourselves:

Thus for me the Other is first the being for whom I am an object; that is, the being *through whom* I gain my objectness. If I am to be able to conceive of even one of my properties in the objective mode, then the Other is already given. [...] In experiencing the look, in experiencing myself as an unrevealed object-ness, I experience the inapprehensible subjectivity of the Other directly and with my being. (Sartre 1943, p. 294)

Accordingly, even our presence to ourselves as a temporal-spatial object in the world is a phenomenon that is mediated by the presence of the other (cf. Sartre 1943, p. 290-291). Moreover, Husserl claims that the same holds for the categories of immanence, appearance, and inwardness. It is only when I experience myself as an object under the gaze of another subject, that I can distinguish my personal inwardness from its public manifestation. Similarly, only when a subject experiences that the same object can be experienced by several subjects, and that it is given for them in various perspectives, that the subject is in a position to realize that there is a distinction between the object itself and its appearance, its being-for-me (cf. Zahavi 1996, pp. 38-39).

What might lived experience be like for a subject who has not been able to perceive others as others? While this is extremely difficult to imagine from our socialized and enculturated perspective, the transformative power of the other is still evident from within the perspective of our own adult existence, so we can entertain some tentative comparative reflections. Merleau-Ponty, for instance, observes that “no sooner has my gaze fallen upon a living body in the process of acting than the objects surrounding it immediately take on a fresh layer of significance: they are no longer simply what I myself

could make of them, they are what this other pattern of behavior is about to make of them” (1945, p. 411-412). This is a good example of how our own sense-making can be shaped by the presence of another subject. Another fitting illustration is Sartre’s description of the experience of being in a public park, facing a lawn with a row of benches along its edge, when a man happens to walk by those benches:

Thus suddenly an object has appeared which has stolen the world from me. Everything is in place; everything still exists for me; but everything is traversed by an invisible flight and fixed in the direction of the new object. The appearance of the Other in the world corresponds therefore to a fixed sliding of the whole universe, to a decentralization of the world which undermines the centralization which I am simultaneously effecting. (Sartre 1943, p. 279)

It is in situations like these, namely when we are prompted to make sense of the world in relation to the perspective of another subject, that we also become aware of our own contributions to the structure of our experience, for example the centralization which we ourselves continually effect on our perceptual world. In this way we can reaffirm that the de-centered presence, which is characteristic of our adult human existence in the world, is co-constituted by the presence of the other.

This completes the brief review of the phenomenology of intersubjectivity. Of special interest was how our experience of self and world is shaped and co-constituted by the possible and actual presence of other subjects. We have seen that this guiding question touched upon the basic perceptual presence of objects, as well as on the conditions for notions of objectivity, appearance, inwardness, and transcendence. Our detached and de-centered presence in the world, which is a condition of possibility for the scientific attitude, turns out to be an intersubjective achievement. On this basis we can now return to the life-mind continuity thesis and provide some clarification of its key concepts.

4. Discussion

The main motivation for this paper was an analysis of the constitutive role of sociality for mind and cognition in order to support the strong life-mind continuity thesis as a unifying framework for cognitive science. We have pursued this goal by drawing on insights developed in two traditions, namely from enactive cognitive science and Husserlian phenomenology. The aim of this final section is to combine these insights in a mutually informative manner, and to reformulate the continuity thesis accordingly.

It should be clear that relating the enactive and the phenomenological traditions in a fruitful manner around the topic of sociality and life-mind continuity is both a challenge and an opportunity. We have argued that a purely systemic (or behavioral) approach to the continuity thesis is not enough, and that we need to incorporate phenomenological considerations. However, this appeal to the observations gained from phenomenological research perspective might cause some resistance in cognitive science, while introducing the enactive approach into the current debates in phenomenology could also be met with some skepticism (cf. Zahavi 2005, p. 176). In order to resolve these tensions it is helpful

to remind ourselves that we are dealing with a subjectivity that is both embodied and embedded, two aspects that can be explored in a complementary fashion by science and phenomenology. And the same applies to intersubjectivity as well. As Merleau-Ponty emphasizes: “we must consider the relation with others *not only as one of the contents of our experience but as an actual structure in its own right*” (Merleau-Ponty 1960, p. 140). Indeed, the enactive approach to social cognition is well positioned to provide the phenomenological tradition with a dynamical account of the interaction process, which might shape the choice of which aspects of experience we turn our attention to and how, while the latter can sharpen the sensitivity of the former to the phenomena that need explaining. Moreover, both traditions are joined in their focus on the primacy of embodied (rather than linguistic) interactions (cf. Zahavi 2005, p. 176). It is with respect to the constitutive impact of such embodied interactions that the enactive approach can be of help to phenomenology, for example by relating the changing qualitative presence of the other to the particular dynamics of changes in ongoing bodily coordination:

[The other’s] autonomy demands frequent readjustments of my individual sense-making. When interaction and individual intentions coordinate, we feel mutually skilful to navigate the interaction: we experience a kind of transparency of the other-in-interaction. But when, for a variety of reasons, a breakdown occurs, and until a new coordination is attained, we experience the other as opaque. (De Jaegher & Di Paolo 2007, p. 504)

It is therefore likely that an integrative methodology that combines both dynamical and phenomenological insights would be mutually beneficial (see also Fuchs and De Jaegher, this issue). For example, one of the central observations regarding intersubjectivity is the immediate presence of others as others; their movements are not mere physical changes but experienced as inherently expressive (even before we get into gesture and language). Enactive cognitive science can explain this phenomenon; in fact, it would expect this to be the case. The meaning that is enacted by our sense-making activity is a *relational* phenomenon (cf. Section 2.2) and is therefore not purely internal to ourselves. On the contrary, the meaning is realized and expressed in the movement of embodied action as such, and is thus accessible to others (unless it is so alien to us that we cannot make sense of what is going on, or is covered up by some kind of camouflage). Note that this conception of sense-making also bridges the traditional gap between instrumental movement and signifying gestures, since both kinds of action are conceived as inherently expressive. The question then becomes in what way the expression of gestures is qualitatively different from that of instrumental movement (relative difference in meaning), thereby avoiding the traditional problem of how it is that gestures turn supposedly meaningless movement into something that expresses some hidden significance (absolute difference in meaning). More work clearly needs to be done to spell this out in more detail; this brief analysis was just intended to indicate of how phenomenology and the enactive approach can develop their insights together.

However, there still remains some tension in relation to the life-mind continuity thesis that must be resolved. The phenomenological analysis of intersubjectivity has led us to argue that “intersubjectivity exists and develops in relation between world-related

subjects, and the world is brought to articulation only in the relation between subjects” (Zahavi 2005, p. 177). However, this stands in contrast to the basic enactive account of agency and sense-making, which posits the bringing forth of a world of significance for the adaptive agent without any mention of the constitutive role of other agents. We will use this lingering tension to our advantage by forcing us to become clearer about the enactive approach to agency and sense-making, as well as the phenomenological notion of worldhood. As indicated in Section 3.1, this tension is not fatal to the enactive account of these basic notions, but it shows that we must be careful that we do not over-anthropomorphize them. But how should we conceive of the world of an agent who is incapable of interacting with others *as* others in their own right?

Jonas claims that “inwardness is coextensive with life” (Jonas 1966, p. 58), a position which has been influential in the development of the notion of sense-making (cf. Weber & Varela 2002; Di Paolo 2005; Thompson 2007). However, on the view presented in this paper, it looks like this is no longer a precise enough description of the phenomenon of life. As Husserl remarks in the case of an imagined solitary human:

For the human being who has not undergone the experience of empathy, or from the standpoint of the abstraction from any empathy, there is no “inwardness” of an “externality”; such a human being would have all of the lived experiences – and all of the objectivities, of whatever sort – that are included under the title of inwardness, but the concept of inwardness would be lost. (Hua XIII/420; quoted by Zahavi 1996, p. 39)

The solitary agent is thus best described as experiencing ‘inwardness’, but an inwardness which is not itself experienced *as* an inwardness. This agent would still be a center of needs and concerns embedded in a meaningful context related to its particular circumstances and viability constraints, as described by Jonas, but this differentiation *as* a center for a world is not a structure of its experience as such (cf. Heidegger 1929). What we have in this solitary case is the ‘organism-*Umwelt*’ dyad that is so well described in the work of the biologist von Uexküll (1934). But in order for the distinction between ‘inner’ and ‘outer’ to become present in experience as such, it is necessary that we are dealing with an intersubjectively constituted form of life, one which unfortunately does not get addressed in the bio-phenomenology of Jonas. Indeed, it is only with a certain process of socialization that there is a possibility of making sense of one’s existence *as* a sense-making existence. Moreover, if we want to talk about the kind of sense-making activities that are involved in enacting the physical world which we experience from our detached human perspective, then appealing to a simple ‘organism-*Umwelt*’ dyad, as would have been done according to the traditional life-mind continuity thesis, is not sufficient. Following the phenomenological tradition, we first have to elucidate the human condition in terms of a ‘self-world-other’ triad. This is because the subject-object dichotomy, which is at the heart of the theoretical attitude that makes abstract knowledge possible, is itself founded on a de-centered and derivative perspective that requires intersubjectivity as its necessary foundation, as we have seen.

The central question, therefore, that needs to be addressed by the life-mind continuity thesis is not how to get from the basic ‘organism-*Umwelt*’ to the human ‘self-world’ structure, but rather how to get from the former to a ‘self-world-*other*’ structure. And then, only on the basis of this transition, will it be possible to determine the conditions for the emergence of the subject-object dichotomy, which has been mistakenly taken as the primary epistemic attitude by cognitive science. Nothing specific has been said about the particular conditions for this dichotomy in this paper, but there is no reason to believe that the life-mind continuity thesis cannot also accommodate this final transition toward detached cognition. Indeed, it appears that Heidegger’s (1927) claim that this transition is caused by break-downs in ongoing coping can be addressed in the framework of enactive cognitive science (cf. Di Paolo & Iizuka 2008).

Finally, one of the challenges for future research will be to better determine the necessary and sufficient conditions which turn the constitutive role of interacting with an other, as described by operational accounts, into the constitutive presence of an other *as* other, as described by phenomenological accounts. In this paper we have only outlined the very beginnings of this project by emphasizing that the phenomenology of intersubjectivity can be complemented by a study of its objective manifestations in the world in terms of interaction dynamics. So far the latter research aspect has revealed that an agent engaging in interactions with others can be constitutively influenced by them even when it is not aware of them as *social* interactions (an insight that is difficult to confirm from the first-person perspective alone). The next step will be to determine the conditions under which such an experiential shift toward a social significance does occur, perhaps through further minimalist psychological studies of inter-individual interactions. Such a project could build on the work by Auvray and colleagues (2009), though it would need to incorporate a principled way of analyzing the lived experience of participants, perhaps by means of a second-person approach to describing experience (e.g. Petitmengin 2006).

5. Conclusion

If enactive cognitive science wants its basic notions of autonomous agency and sense-making to be the foundation for a unified theory of life and mind, then it needs to appeal to a strong version of the life-mind continuity thesis. Only if this thesis turns out to be correct can this approach to cognitive science systematically reject the criticism that it is dealing with somewhat interesting, but ultimately hardly relevant, descriptions of low-level ‘implementation’ details. So far, however, proponents of the life-mind continuity thesis have been plagued by what we have called the cognitive gap, namely an inability to conceive of how the biological principles applicable to simple organisms can be used to explain the highest reaches of human cognition.

Following De Jaegher and Froese (in press), we have argued that this conceptual impasse is largely due to the methodological individualism that is prevalent in cognitive science, and that a proper consideration of the constitutive role of sociality for agency and sense-making is needed in order to make the thesis plausible. More precisely, we have proposed that what is needed is a change of several traditional premises: (i) internalism must be replaced by a distributed view of cognition, (ii) concern for individual agency must be

complemented with a systematic consideration of the constitutive role of sociality, (iii) an individual's engagement with social elements must be understood as primarily active rather than passive, and (iv) these social elements must primarily be viewed as dynamic, emergent processes resulting out of interactions, rather than static, physical structures that are pre-given. De Jaegher and Froese (in press) have developed these changes from a systemic perspective; here we have complemented these efforts by introducing some supporting phenomenological insights. We suggest on the basis of these considerations that the strong life-mind continuity thesis advocated by the enactive cognitive science is a promising contender for a unifying thesis of recent cognitive science research.

References

- Auvray, M., Hanneton, S., Lenay, C. & O'Regan, K. (2005), "There is something out there: Distal attribution in sensory substitution, twenty years later", *Journal of Integrative Neuroscience*, **4**(4), pp. 505-521
- Auvray, M., Lenay, C. & Stewart, J. (2009), "Perceptual interactions in a minimalist virtual environment", *New Ideas in Psychology*, **27**(1), pp. 32-47
- Barandiaran, X., Di Paolo, E. A., Rohde, M. (forthcoming), "Defining agency: Individuality, normativity, asymmetry and spatio-temporality in action", *Adaptive Behavior*
- Beer, R. D. (2000), "Dynamical approaches to cognitive science", *Trends in Cognitive Sciences*, **4**(3), pp. 91-99
- Boden, M. A. (2006), "Of Islands and Interactions", *Journal of Consciousness Studies*, **13**(5), pp. 53-63
- Clark, A. (2001), *Mindware: An Introduction to the Philosophy of Cognitive Science*, Oxford, UK: Oxford University Press
- Clark, A. (2003), *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*, Oxford, UK: Oxford University Press
- Clark, A. (2005), "Beyond the Flesh: Some Lessons from a Mole Cricket", *Artificial Life*, **11**(1-2), pp. 233-244
- Clark, A. (2008), *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*, New York, NY: Oxford University Press
- De Jaegher, H. & Di Paolo, E. A. (2007), "Participatory sense-making: An enactive approach to social cognition", *Phenomenology and the Cognitive Sciences*, **6**(4), pp. 485-507
- De Jaegher, H. & Froese, T. (in press), "On the role of social interaction in individual agency", *Adaptive Behavior*, forthcoming
- Di Paolo, E. A. (2005), "Autopoiesis, adaptivity, teleology, agency", *Phenomenology and the Cognitive Sciences*, **4**(4), pp. 429-452
- Di Paolo, E. A. (2009), "Extended Life", *Topoi*, **28**(1), pp. 9-21
- Di Paolo, E. A., Rohde, M. & De Jaegher, H. (in press), "Horizons for the Enactive Mind: Values, Social Interaction, and Play", in: J. Stewart, O. Gapenne & E. A. Di Paolo (eds.), *Enaction: Towards a New Paradigm for Cognitive Science*, Cambridge, MA: The MIT Press, forthcoming
- 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), pp. 278-294
- Donaldson, M. (1992) *Human minds: An exploration*, London, UK: Penguin Books

- Froese, T. (2009), "Hume and the enactive approach to mind", *Phenomenology and the Cognitive Sciences*, **8**(1), pp. 95-133
- Froese, T. & Di Paolo, E. A. (2008), "Stability of coordination requires mutuality of interaction in a model of embodied agents", in: M. Asada, *et al.* (eds.), *From Animals to Animats 10: Proc. of the 10th Int. Conf. on Simulation of Adaptive Behavior*, Berlin, Germany: Springer-Verlag, pp. 52-61
- Froese, T. & Di Paolo, E. A. (in press), "Modeling social interaction as perceptual crossing: An investigation into the dynamics of the interaction process", *Connection Science*, forthcoming
- Froese, T., Virgo, N. & Izquierdo, E. (2007), "Autonomy: a review and a reappraisal", in: F. Almeida e Costa *et al.* (eds.), *Advances in Artificial Life: Proc. of the 9th Euro. Conf. on Artificial Life*, Berlin, Germany: Springer-Verlag, pp. 455-464
- Froese, T. & Ziemke, T. (2009), "Enactive Artificial Intelligence: Investigating the systemic organization of life and mind", *Artificial Intelligence*, **173**(3-4), pp. 366-500
- Gallagher, S. (2001), "The practice of mind: Theory, simulation, or primary interaction", *Journal of Consciousness Studies*, **8**(5-7), pp. 83-108
- Gallagher, S. (2008a), "Intersubjectivity in perception", *Continental Philosophy Review*, **41**(2), pp. 163-178
- Gallagher, S. (2008b), "Direct perception in the intersubjective context", *Consciousness and Cognition*, **17**(2), pp. 535-543
- Gallagher, S. & Zahavi, D. (2008), *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science*, London, UK: Routledge
- Godfrey-Smith, P. (1996), "Spencer and Dewey on Life and Mind", in: M. Boden (ed.), *The Philosophy of Artificial Life*, Oxford, UK: Oxford University Press, pp. 314-331
- Heidegger, M. (1927), *Sein und Zeit*, trans. by: J. Macquarrie & E. Robinson, *Being and Time*, Oxford, UK: Blackwell Publishing Ltd., 1962
- Heidegger, M. (1929), *Die Grundbegriffe der Metaphysik: Welt, Endlichkeit, Einsamkeit*, trans. by: W. McNeill & N. Walker, *The Fundamental Concepts of Metaphysics: World, Finitude, Solitude*, Bloomington, IN: Indiana University Press, 1995
- Husserliana (Hua)*
- Husserliana I: Husserl, E. (1929), *Cartesianische Meditationen und Pariser Vorträge*, Den Haag, Netherlands: Martinus Nijhoff, 1950; trans. by: D. Cairns, *Cartesian Meditations: An Introduction to Phenomenology*, The Hague, Netherlands: Martinus Nijhoff, 1960
- Husserliana IX: Husserl, E. (1925), *Phänomenologische Psychologie. Vorlesungen Sommersemester 1925*, Den Haag, Netherlands: Martinus Nijhoff, 1962; trans. by: J. Scanlon, *Phenomenological Psychology: Lectures, Summer Semester, 1925*, The Hague, Netherlands: Martinus Nijhoff, 1977
- Husserliana X: Husserl, E. (1893-1917), *Zur Phänomenologie des inneren Zeitbewusstseins (1893-1917)*, Den Haag, Netherlands: Martinus Nijhoff, 1966; trans. by: J. B. Brough, *On the Phenomenology of the Consciousness of Internal Time (1893-1917)*, Dordrecht, Netherlands: Kluwer Academic, 1991
- Husserliana XIII: Husserl, E. (1905-1920), *Zur Phänomenologie der Intersubjektivität. Texte aus dem Nachlass. Erster Teil: 1905-1920*, Den Haag, Netherlands: Martinus Nijhoff, 1973
- Husserliana XIV: Husserl, E. (1921-1928), *Zur Phänomenologie der Intersubjektivität. Texte aus dem Nachlass. Zweiter Teil: 1921-1928*, Den Haag, Netherlands: Martinus Nijhoff, 1973
- Husserliana XV: Husserl, E. (1929-1935), *Zur Phänomenologie der Intersubjektivität. Texte aus dem Nachlass. Dritter Teil: 1929-1935*, Den Haag, Netherlands: Martinus Nijhoff, 1973
- Hutchins, E. (1995), *Cognition in the Wild*, Cambridge, MA: The MIT Press
- Jonas, H. (1966), *The Phenomenon of Life: Toward a Philosophical Biology*, Evanston, IL: Northwestern University Press, 2001

- Jonas, H. (1992), "The Burden and Blessing of Mortality", *The Hastings Center Report*, **22**(1), pp. 34-40
- Kant, I. (1790), *Kritik der Urteilskraft*, trans. by: W. S. Pluhar, *Critique of Judgment*, Indianapolis, IN: Hackett Publishing Company, 1987
- Maturana, H. R., & Varela, F. J. (1987), *The Tree of Knowledge: The Biological Roots of Human Understanding*, Boston, MA: Shambhala Publications
- Merleau-Ponty, M. (1945), *Phénoménologie de la perception*, trans. by: C. Smith, *Phenomenology of perception*, New York, NY: Routledge & Kegan Paul, 1962
- 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.), Evanston, IL: Northwestern University Press, pp. 96-155
- Moreno, A. & Etxeberria, A. (2005), "Agency in Natural and Artificial Systems", *Artificial Life*, **11**(1-2), pp. 161-175
- Mossio, M. & Taraborelli, D. (2008), "Action-dependent perceptual invariants: From ecological to sensorimotor approaches", *Consciousness and Cognition*, **17**(4), pp. 1324-1340
- Noë, A. (2004), *Action in Perception*, Cambridge, MA: The MIT Press
- Petitmengin, C. (2006), "Describing one's subjective experience in the second person: An interview method for the science of consciousness", *Phenomenology and the Cognitive Sciences*, **5**(3-4), pp. 229-269
- Piaget, J. (1967), *Biologie et connaissance*, Paris, France: Editions Gallimard; transl. as: *Biology and Knowledge: An Essay on the Relations between Organic Regulations and Cognitive Processes*, Chicago, IL: The University of Chicago Press, 1971
- Sartre, J.-P. (1943), *L'Être et le néant*, trans. by H. E. Barnes, *Being and Nothingness: An essay on phenomenological ontology*, Oxon, UK: Routledge, 2003
- Steiner, P. & Stewart, J. (in press), "From Autonomy to Heteronomy (and back): The Enaction of Social Life", *Phenomenology and the Cognitive Sciences*, forthcoming
- Stewart, J. (1996), "Cognition = life: Implications for higher-level cognition", *Behavioral Processes*, **35**, pp. 311-326
- Stewart, J. (in press), "Foundational issues in enaction as a paradigm for cognitive science: From the origin of life to consciousness and writing", in: J. Stewart, O. Gapenne & E. A. Di Paolo (eds.), *Enaction: Towards a New Paradigm for Cognitive Science*, Cambridge, MA: The MIT Press, forthcoming
- Thompson, E. (2001), "Empathy and Consciousness", *Journal of Consciousness Studies*, **8**(5-7), pp. 1-32
- Thompson, E. (2004), "Life and mind: From autopoiesis to neurophenomenology. A tribute to Francisco Varela", *Phenomenology and the Cognitive Sciences*, **3**(4), pp. 381-398
- Thompson, E. (2007), *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*, Cambridge, MA: The Belknap Press of Harvard University Press
- Tomasello, M. (1999), *The Cultural Origins of Human Cognition*, Cambridge, MA: Harvard University Press
- Tomasello, M., Call, J. & Hare, B. (2003), "Chimpanzees understand psychological states – the question is which ones and to what extent", *Trends in Cognitive Sciences*, **7**(4), pp. 153-156
- Trevarthen, C. & Hubley, P. (1978), "Secondary intersubjectivity: Confidence, confiding and acts of meaning in the first year", in: A. Lock (ed.), *Action, gesture and symbol: The emergence of language*, London, UK: Academic Press, pp. 183-229
- Varela, F. J. (1997), "Patterns of Life: Intertwining Identity and Cognition", *Brain and Cognition*, **34**(1), pp. 72-87

- 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*, Stanford, CA: Stanford Uni. Press, pp. 266-317
- Varela, F. J., Maturana, H. R. & Uribe, R. (1974), "Autopoiesis: The organization of living systems, its characterization and a model", *BioSystems*, **5**, pp. 187-196
- Varela, F. J., Thompson, E. & Rosch, E. (1991), *The Embodied Mind: Cognitive Science and Human Experience*, Cambridge, MA: The MIT Press
- Von Foerster, H. (1976), "Objects: Tokens for (Eigen-)Behaviors", *ASC Cybernetics Forum*, **8**(3-4), pp. 91-96
- von Uexküll, J. (1934), "Streifzüge durch die Umwelten von Tieren und Menschen: Ein Bilderbuch unsichtbarer Welten", trans. by: C. H. Schiller, "A stroll through the worlds of animals and men: a picture book of invisible worlds", in: C. H. Schiller (ed.), *Instinctive Behavior: The Development of a Modern Concept*, New York, NY: International Universities Press, 1957, pp. 5-80. Also appeared in: *Semiotica*, **89**(4), 1992, pp. 319-391
- Weber, A. & Varela, F. J. (2002), "Life after Kant: Natural purposes and the autopoietic foundations of biological individuality", *Phenomenology and the Cognitive Sciences*, **1**, pp. 97-125
- Wheeler, M. (1997), "Cognition's Coming Home: the Reunion of Life and Mind", in: P. Husbands & I. Harvey (eds.), *Proc. of the 4th Euro. Conf. on Artificial Life*, Cambridge, MA: The MIT Press, pp. 10-19
- Zahavi, D. (1996), *Husserl und die transzendente Intersubjektivität: Eine Antwort auf die sprachpragmatische Kritik*, trans. by: E. A. Behnke, *Husserl and Transcendental Intersubjectivity: A response to the Linguistic-Pragmatic Critique*, Athens, OH: Ohio University Press, 2001
- Zahavi, D. (2001), "Beyond Empathy: Phenomenological Approaches to Intersubjectivity", *Journal of Consciousness Studies*, **8**(5-7), pp. 151-167
- Zahavi, D. (2005), *Subjectivity and Selfhood: Investigating the First-Person Perspective*, Cambridge, MA: The MIT Press
- Zlatev, J. (2002), "Meaning = Life (+ Culture): An outline of a unified biocultural theory of meaning", *Evolution of Communication*, **4**(2), pp. 253-296