The Enactive Approach

Ezequiel Di Paolo and Evan Thompson

Forthcoming in Lawrence Shapiro, ed., *The Routledge Handbook of Embodied Cognition* (Routledge Press).

Embodied approaches to cognition hold that the body is crucial for cognition. Yet despite many decades of research,[1] what this "embodiment thesis"[2] (Wilson and Foglia 2011) amounts to still remains unclear, as the present volume with its diverse range of contributions indicates (see also Shapiro 2011). How to interpret the embodiment thesis depends on how to interpret the meanings of its key terms, "body" and "cognition," as well as on what it means exactly to say that the body is "crucial" for cognition (Kyselo & Di Paolo, under review). In recent years, the term "embodied" has been used elastically to refer to anything from conservative ideas about how bodily action provides a format for neuronal representations (Goldman and de Vignemont 2009; Gallese 2010; Goldman 2012) or helps to reduce computational load (Clark 2008; Wheeler 2005, 2010; Wilson 2004), to a variety of "radical embodiment" (Clark 1999; Thompson and Varela 2001) proposals—for example, that kinesthetic body schemas are a constitutive part of mental skills (Lakoff and Johnson 1999; Nuñez 2010), that sensorimotor know-how is a constitutive part of perceptual experience (O'Regan and Noë 2001; Noë 2004), that bodily life-regulation is a constitutive part of phenomenal consciousness and its extended neurophysiological substrates (Thompson and Varela 2001; Thompson and Cosmelli 2011), and that social sensorimotor interaction can be a constitutive part of social cognition (De Jaegher, Di Paolo, and Gallagher 2010). In some cases, these radical embodiment proposals are based on the "enactive" view that cognition depends constitutively on the living body, understood as an *autonomous* system (Varela, Thompson, and Rosch 1991; Thompson 2007; Di Paolo 2009; Froese and Ziemke 2009). Our aim in this chapter is to explain this key enactive notion of autonomy and why it is needed if embodied cognitive science is to offer a genuine alternative to more traditional functionalist and cognitivist views.

The Body and Autonomy

A key attribute of the living body is its individuation, the process by which it makes itself distinct from its immediate surroundings and that enables an observer to distinguish it as an identifiable entity. More precisely, a key attribute of the body is that it is *self-individuating*—it generates and maintains itself through constant structural and functional change. Yet many of the systems we study in science—particles, rivers, communities, galaxies, and even bodies—we typically individuate from the outside by convention, with varying degrees of accuracy. In other words, what counts as one system versus another typically depends on the conventional criteria we use to individuate the system; such criteria include considerations of convenience, perceptual biases, longer versus shorter relative timescales of change, semi-arbitrary definitions, or historical or practical use. To this day, functionalism in cognitive science identifies cognitive systems in precisely this way, that is, by convention: A robot, a body, or an agent is specified as the system it is according to some tacit convention of perception, use, or measurement. Most embodied approaches to the mind also proceed in this way. Thus references to the "body" are understood contextually as references to a given anatomical structure or physiological function, to the details of a given sensorimotor system, or to being situated in a given world of habits, norms, skills, and so on. Yet, regardless of whatever conventional criteria we may use to individuate the body, the body—like all living systems—nonetheless possesses the peculiar attribute of being self-individuating, such that it actively generates and maintains the distinction between itself and its environment. A guiding idea of the enactive approach is that any adequate account of how the body can either be or instantiate a cognitive system must take account of this fact that the body is self-individuating.

This point brings us to the principal concept that differentiates enactivism from other embodied approaches to the mind—the concept of *autonomy*. By making use of this concept—which, as we will see, explains how bodies are self-individuating—we can give operational criteria for distinguishing cognitive systems from non-cognitive systems.

The idea of autonomy has roots in Maturana and Varela's (1980) theory of autopoiesis. This theory has strongly influenced enactive thinking to the point that sometimes the style of enactivism based on the notion of autonomy has been branded "autopoietic enactivism" (Hutto and Myin, 2013). This easy label is descriptive but potentially misleading, for the key theoretical developments in this area concern various forms of extensions—perhaps even departures—from the original autopoietic framework (see Di Paolo 2009; Thompson 2011a, 2011b). These elaborations should become clear in this section for the case of autonomy (the same goes for other concepts in the enactive framework not discussed here in full, such as adaptivity, sense-making, and agency; see Di Paolo 2005, 2009, Thompson 2007, 2011a, 2001b).

Autonomy was initially conceived as a generalization of the concept of autopoiesis or self-production (Varela 1979). The concept of autopoiesis describes a peculiar aspect of the organization of living organisms, namely, that their ongoing processes of material and energetic exchanges with the world, and of internal transformation and metabolizing, relate to each other in such a way that the same organization is constantly regenerated by the activities of the processes themselves, despite whatever variations occur from case to case. Varela (1979) extended this idea into other domains. Thus, he identified a similar organizational logic in the animal nervous system and immune networks, and he hinted at the application of the concept of autonomy to other domains, such as communication networks and conversations.

An autonomous system is defined as an *operationally-closed* and *precarious* system. Before unpacking these terms, it may help to look at the intuitive procedure that we can apply to tell whether or not we are observing an autonomous system.

Consider the following situation. A scientist is observing, recording and intervening in various processes that she finds of interest. Either in the lab or in the field, she observes and measures variables and their relations, takes notes of events, and attempts to establish the connections between her observations. To a large extent, *what* she chooses to observe and record is a matter of choice, as are the procedures she follows. She may be interested in the fact that the temperature in the lab seems relatively constant throughout the day in spite of the varying temperatures outside. Or she may be looking at some chemical reactions and measuring the speed with which they happen. To keep to a general level, let us say that she is observing *processes*. How each of these processes is identified is not important here; the means of identification will likely depend on the observer's history, skills, tools, and purposes.

As the scientist intervenes in the various processes she is observing, she takes note of regular effects on other processes. For example, she changes a setting in a thermostat and the room temperature is now maintained at a higher constant level. Or she observes that the chemical reactions are now happening at a higher rate, and so on. Eventually, she may be able to establish various relations between the observed processes. Some of these relations indicate merely contextual effects—the observed process is altered in some way by intervening on a different process—other relations are stronger and indicate enabling conditions—the observed process disappears or stops if a particular intervention is made on a different process (De Jaegher, Di Paolo, and Gallagher 2010). Thus, modifying the thermostat setting brings the desired room temperature to a cooler or warmer value, but unplugging the air conditioning simply prevents any temperature regulation from happening at all. Temperature regulation as such has stopped as a process. Such enabling relations, of course, depend on the circumstances.

We can use Figure 1 to depict this kind of situation. The circles represent the processes being observed by the scientist. Whenever an enabling relation is established, the scientist draws an arrow going from the process that is perturbed to the process that stops or disappears as a consequence. An arrow going from process A to process B indicates that A is an enabling condition for B to occur. Of course, there may be several enabling conditions. We do not assume that the scientist is mapping all of them, only those that she finds relevant or can uncover with her methods.

As the mapping of enabling conditions proceeds, the scientist makes an interesting observation. There seems to be a set of processes that relate to each other with a special topological property. These processes are marked in black in the figure. If we look at any process in black, we observe that it has some enabling arrows *arriving at it* that originate in other processes in black, and moreover, that it has some enabling arrows *coming out of it* that end up also in other processes in black. When this condition is met, the black processes form a network of enabling relations; this network property is what we mean by *operational closure*.

Notice that this form of "closure" does not imply the independence of the network on other processes that are not part of it. Firstly, there may be enabling dependencies on external processes that are not themselves enabled by the network; for example, plants can photosynthesize only in the presence of sunlight (an enabling condition), but the sun's existence is independent of plant life on earth. Similarly, there may be processes that are made possible only by the activity of the network but that do not themselves "feed back" any enabling relations toward the network itself. Secondly, the arrows do not describe any possible form of coupling between processes, but rather only a link of enabling dependence. Other links may exist too, such as interactions that have only contextual effects. In short, an operationally closed system should not be conceived as isolated from dependencies or from interactions.

Notice too that although the choice of processes under study is more or less arbitrary and subject to the observer's history, goals, tools, and methods, the topological property unraveled is not arbitrary. The operationally closed network could be larger than originally thought, as new relations of enabling dependencies are discovered. But it is already an operationally closed network by itself and this fact cannot be changed short of its inner enabling conditions changing, that is, short of some of its inner processes stopping.

A living cell is an example of an operationally closed network. The closed dependencies between constituent chemical and physical processes in the cell are very complex, but it is relatively easy to see some of them. For example, the spatial enclosure provided by a semi-permeable cell membrane is an enabling condition for certain autocatalytic reactions to occur in the cell's interior, otherwise the catalysts would diffuse in space and the reactions would occur at a much different rate or not at all. Hence there is an enabling arrow going from the spatial configuration of the membrane to the metabolic reactions. But the membrane containment is not a given; the membrane is also a precarious process that depends, among other things, on the repair components that are generated by the cell's metabolism. So there is an enabling arrow coming back from metabolic reactions to the membrane. Hence we have already identified an operationally closed loop between these cellular processes. If the scientist chose not to observe the membrane in relation to the metabolic reactions, she would probably miss the topological relation between them. Operational closure—cellular life in this case—can be missed if we choose to put the focus of observation elsewhere, but it is not an arbitrary property if we observe it at the right level.

Is autonomy the same as operational closure? Although it seems to have been the same for Varela (1979)—at least in terms of his formal definition of an autonomous system—we have argued in other writings that operational closure is not sufficient to capture certain important implications of the wider sense of the term "autonomy"—implications conveyed by notions such as "spontaneity," "immanent purposiveness," "intrinsic teleology," and the "self-generation of norms" (see Di Paolo 2005; Thompson 2007). Given the above definition of operational closure, various trivial examples of such closure may exist. For example, in cellular automata, the regeneration of an equilibrium state in each cell mutually depends on the equilibrium state in others, making the dependencies into a closed network.

We need an additional condition to make operational closure non-trivial, and this condition is that of *precariousness*. Of course, all material processes are precarious if we wait long enough. In the current context, however, what we mean by "precariousness" is the following condition: In the absence of the enabling relations established by the operationally closed network, a process belonging to the network will stop or run down.

It might seem at first that this condition is redundant. Why is precariousness not implied in the fact that a black circle in Figure 1 is always enabled by other black circles? So, surely if the enabling relations disappear, that process should stop. This stoppage would be the case if the scientist had exhaustively found all the enabling relations, but this need not be the case. There may also be redundancies among the enabling relations affecting a given process (more than one arrow may come into a circle, and this fact could mean either that all of the arrows are needed simultaneously or that different subsets of the incoming arrows are sufficient for the process to continue according to the conditions). An enabling relation, as we stated above, in principle holds only in given circumstances. When a process is enabled by the operationally closed network and by external processes as well, if the network is removed and the process remains—in the new circumstances—thanks only to the external support, then that process is not precarious. A precarious process is such that, whatever the complex configuration of enabling conditions, if the dependencies on the operationally closed network are removed, the process necessarily stops. In other words, it is not possible for a precarious process in an operationally closed network to exist on its own in the circumstances created by the absence of the network.

A precarious, operationally closed system is literally self-enabling, and thus it sustains itself in time partially due to the activity of its own constituent processes. Moreover, because these processes are precarious, the system is always decaying. The "natural tendency" for each constituent process is to stop, a fate the activity of the other processes prevents. The network is constructed on a double negation. The impermanence of each individual process tends to affect the network negatively if sustained unchecked for a sufficient time. It is only the effect of other processes that curb these negative tendencies. This dynamic contrasts with the way we typically conceive of organic processes as contributing positively to sustaining life; if any of these processes were to run unchecked, it would kill the organism. Thus, a precarious, operationally closed system is inherently restless, and in order to sustain its intrinsic tendencies towards internal imbalance, it requires energy, matter, and relations with the outside world. Hence, the system is not only self-enabling, but also shows spontaneity in its interactions due to a constitutive need to constantly "buy time" against the negative tendencies of its own parts.

The simultaneous requirement of operational closure and precariousness are the defining properties of autonomy for enactivism. It is this concept of autonomy that answers the opening question in this section about the individuation of the body. A body is understood as such an autonomous system, an understanding that allows for the possibility that any given body need not be constituted exclusively by its biochemical or physiological processes (Thompson and Stapleton 2009; Kyselo and Di Paolo, under review).

Notice that what we have provided in this section is a clear, step-by-step procedure to answer empirically the question of whether a system is autonomous. The enactive concept of autonomy is entirely operational, and therefore naturalistic, though not reductionist. It is, however, not a concept that could be captured using the methods of functionalism, and this point already entails a fundamental difference between enactive and functionalist approaches to the mind. As we have seen, autonomy requires precariousness, but precariousness is not a positive property of a process, but rather an unavoidable aspect of materiality. In the current context, precariousness is the insufficient permanence of any positive property that might play a positive functional role in sustaining the autonomous system in the absence of the system's closed organization. Precariousness cannot be "revealed" as a positive property and yet its negative effects are what the system is constantly acting against. For this reason, precariousness might be partially modeled in functionalistic terms, but never fully captured in those terms, as the conditions that satisfy any functional approximation (for instance, a decay function affecting various parameters) are in fact themselves precarious if the system is really autonomous.

In this way, the enactive concept of autonomy captures individuation as a non-arbitrary ongoing process as well as the spontaneity of living bodies. The concept also leads to various other developments, such as a naturalistic way to account for what Kant described as the intrinsic teleology or immanent purposiveness that appears to belong to living beings by virtue of being self-organizing (Weber and Varela 2002; Thompson 2007).

Linked to the notion of autonomy are two other key enactive notions, those of *adaptivity* and *sense-making* (Di Paolo 2005, 2009; Thompson 2011a, 2001b). "Adaptivity" refers to the ability of certain autonomous systems to regulate their operationally closed processes in relation to conditions registered as improving or deteriorating, viable or unviable (Di Paolo 2005). This ability for adaptive regulation is inextricably linked to autonomy insofar as it happens with respect to the implications for the continuation of the system's autonomous identity. "Sense-making" describes behaviour or conduct in relation to norms of interaction that the system itself brings forth on the basis of its adaptive autonomy. An adaptive autonomous system produces and sustains its own identity in precarious conditions, registered as better or worse, and thereby establishes a perspective from which interactions with the world acquire a normative status. Certain interactions facilitate autonomy and other interactions degrade it. In Merleau-Ponty's words: ''each organism, in the presence of a given milieu, has its optimal conditions of activity and its proper manner of realizing equilibrium,'' and each organism ''modifies its milieu according to the internal norms of its activity'' (Merleau-Ponty 1963, pp. 148, 154). For the enactive approach, a system is cognitive when its behaviour is governed by the norm of the system's own continued existence and flourishing. Basic cognition, on this view, is not a matter of representing states of affairs but rather of establishing relevance through the need to maintain an identity that is constantly facing the possibility of disintegration. From this perspective, the body is not just the means but also an end of being a cognitive system. To put the point another way, basic cognition is more a matter of adaptive self-regulation in precarious conditions than abstract problem solving. The point here is not to deny that we can and do engage in high-level problem solving. Rather, it is to say that this kind of narrow cognition presupposes the broader and more basic cognition that we call sense-making.

Applications of Enactive Ideas

Modelling autonomy; bacterial chemotaxis

The definition of autonomy we have given in the previous section aims at providing a procedure through which the question of whether a specific system is autonomous or not can be settled. It should then be possible to construct models of autonomous systems that enable us to explore some of the implications of this concept.

We begin with a couple of clarifications. First, by a "model" we mean a formal or material system that instantiates some of the properties of an autonomous system that are relevant in a particular context. By its nature, a model is a simplification and not a realization of the system being modelled. In other words, a model of an autonomous system is not, in itself, an autonomous system, nor is it meant to be, in the same sense that a fluid dynamics model of a river is not a river. The second clarification concerns the question of whether, given some modelling techniques, it is possible to capture all of the aspects of autonomy. This is an open question, but we can at least say that given our comments on precariousness and the fact that precariousness cannot be captured as a positive function, it is not possible to model full autonomy in traditional computational terms, including dynamical systems models. It may be the case that a full model of autonomy requires a material system that exhibits the relations of precarious operational closure among the processes that are involved in that system. A computer does not typically work in this way. Nevertheless, in the context of the goal of a computational model, we may approximate with sufficient accuracy those aspects that are relevant to the question at hand. This way of proceeding is no different from other kinds of scientific models; such models have a limited range of validity outside of which the modelling assumptions fail.

Matthew Egbert and colleagues have explored a series of dynamical systems models of autonomy in the context of bacterial behaviour (Egbert et al. 2010). These models explore the relation between autonomy and sense-making in its most basic forms, and so can help fine-tune enactive theory (bacterial chemotaxis is a canonical example of sense-making, often used by Varela himself; see Varela 1991; Thompson 2007; 2011c). Egbert and colleagues take note of an interesting situation in bacterial research. For decades, the chemotactic behaviour of bacteria has been investigated as if it did not depend on their metabolic state. Whether bacteria are "hungry" or "satiated," it is assumed that they will always follow a chemical gradient towards sources of nourishment. Evidence shows, however, that the assumption of metabolism-independence often fails (Alexandre and Zhulin 2001). For example, in some species nonmetabolizable chemical analogues of metabolizable attractants are not themselves attractants (even though they excite chemical sensors in the same way). In addition, the inhibition of the metabolism of a chemical attractant completely stops chemotaxis to this attractant and only this attractant. Furthermore, the presence of another metabolizable chemical prevents chemotaxis to all attractants studied. In short, there seems to be a deep link between metabolism, the autonomous process of self-construction, and behaviour—the regulation of interactions with the world—as expected from the relation between autonomy and sense-making proposed by enactivism. By modelling metabolism as a cycle of autocatalytic reactions far from thermodynamic equilibrium, thus capturing some aspects of precarious operational closure, and by modelling behaviour regulation as modulated by metabolism, Egbert and colleagues have managed to replicate various empirically observed behaviours of bacteria. These behaviours are chemotaxis towards metabolic resources and away from metabolic inhibitors, inhibition of chemotaxis in the presence of abundant resources, cessation of chemotaxis to a resource due to inhibition of the metabolism of that resource, sensitivity to metabolic and behavioural history, and integration of simultaneous complex environmental ''stimuli." Various extensions of this model make it possible to explore the evolutionary implications of this basic link between life and mind proposed by enactivism (Egbert et al. 2012; Egbert 2013; Barandiaran and Egbert 2013).

Participatory sense-making

Another area where enactive ideas are making an impact is intersubjectivity research. This area has been a concern of enactivism for some time (Thompson 2001, 2007). As a result of De Jaegher and Di Paolo's (2007) introduction of the concept of "participatory sense-making," enactivism has begun to ramify into fields as varied as animal social behaviour, psychiatry and psychopathology, social neuroscience, dance, music, literature and education studies, and embodied linguistics (see De Jaegher, Di Paolo, and Gallagher 2010).

The enactive account of intersubjectivity calls attention to participatory and non-individualistic processes, unlike traditional approaches where social understanding is reduced to the inferences or simulations a typically passive observer can make about the intentions of others based on their external behaviour. Making explicit the domain of social interaction is to take a crucial step away from methodological individualism. For this step, De Jaegher and Di Paolo (2007) use the idea of autonomy, but now applied to the relational processes that often take place in the encounter of two or more people. Accordingly, sometimes these encounters take on a life of their own and the actions of the agents involved, or their intentions, do not fully determine the outcome of the encounter, which also depends on its own relational and dynamic constituent processes. In these cases, where relational patterns emerge as autonomous and the interactors themselves remain autonomous, we are in the presence of a social interaction, which can be defined formally (De Jaegher and Di Paolo 2007, p 493). The processes involved are patterns of intercorporeal coordination at various levels—imitative gestures, regulation of personal distance, posture and orientation, attunement to conversation or walking rhythms, and so on (such patterns have long been investigated in psychology and social science).

Given that sense-making is an embodied process of active regulation of the coupling between agent and world, social interaction—through patterns of bodily coordination and breakdown—opens the possibility of this process being shared among the interactors. This shared form of sense-making is what is meant by "participatory sense-making." It happens to various degrees, from orientation of individual sense-making (someone draws our attention to an aspect of the world we have ignored) to joint sense-making (a piece of work is literally created together through a process that would not be possible by the individuals involved on their own).

This proposal has important empirical implications. For example, De Jaegher, Di Paolo, and Gallagher (2010) show that making the scientific study of interaction more explicit can offer new hypotheses about the processes sustaining a cognitive performance, and frees researchers from postulating complicated mechanisms in the individual brain that duplicate what the interactive dynamic configuration is able to bring about on its own. This proposal of paying more attention to interactive factors comes at a time when empirical methods in psychology and neuroscience (such as hyperscanning, thermal imaging, motion energy analysis, and second-person methodologies) have started to develop to the point that live and relatively unconstrained interactions can be directly investigated. These sciences may now move away from the traditional individualistic and observational paradigms. It is even possible to propose broad empirical hypotheses about the relation between individual brain mechanisms, the development of interactive skills, and interactive history (Di Paolo and De Jaegher 2012).

Through participatory sense-making, the enactive approach thematizes pre-existing empirical and practical knowledge that has often been neglected by mainstream theoretical frameworks. Such thematization is particularly the case in psychiatry, where the social dimension in the aetiology, diagnosis, and intervention of disorders such as schizophrenia or autism has been well known in practice and to a large extent well documented, and yet cognitivist or neurobiological approaches have tended to downplay this dimension. Explanations of mental disorders too can be reconsidered from a non-individualistic enactive perspective (e.g., de Haan and Fuchs 2010; De Jaegher, 2013).

Such reconsideration does not imply positing either the individual or the interactive levels as fundamental, but rather understanding the mutually enabling relations between the two levels (Fuchs 2012; De Jaegher & Di Paolo 2013).

Conclusion

In conclusion, we can use the concepts sketched here to present brief enactive responses to the following three questions about the embodiment thesis that the body is crucial for cognition: (1) What is meant by *body*? (2) What is meant by *cognition*? (3) What is meant by *crucial*?

(1) We have seen that enactivism, unlike other approaches, attempts to provide a principled definition of the body as a self-individuating system. The concept of autonomy is what allows us to provide this definition. Thus, what is meant by "body," for the enactive approach, is not the body as a functional system defined in terms of inputs and outputs—as it is for functionalist cognitive science—but rather the body as an adaptively autonomous and therefore sense-making system.[3]

(2) Cognition, in its most general form, is sense-making—the adaptive regulation of states and interactions by an agent with respect to the consequences for the agent's own viability.

(3) Without a body, there cannot be sense-making. Moreover, sense-making is a bodily process of adaptive self-regulation. The link between the body and cognition is accordingly constitutive and not merely causal. To be a sense-maker is, among other things, to be autonomous and precarious, that is, is to be a body, in the precise sense of "body" that the enactive approach indicates.

Further Reading

The "enactive approach" to cognition was first proposed by Francisco Varela, Evan Thompson, and Eleanor Rosch, *The Embodied Mind: Cognitive Science and Human Experience* (Cambridge, MA: The MIT Press, 1991). The concept of "autonomy" (defined in terms of operational closure) was introduced by Francisco Varela, *Principles of Biological Autonomy* (New York: Elsevier, North Holland, 1979). For an extensive contemporary statement of the enactive approach, covering many new developments in cognitive science, theoretical biology, and philosophy, see Evan Thompson, *Mind in Life: Biology, Phenomenology, and the Sciences of Mind* (Cambridge, MA: Harvard University Press, 2007), as well as the "*Précis* of Mind in Life," together with "Commentaries" and the "Author's Reply," published by the *Journal of Consciousness Studies*, vol. 18, 2011. For the implications enactive approach for AI, see Tom Froese and Tom Ziemke, "Enactive Artificial Intelligence: Investigating the Systemic Organization of Life and Mind," *Artificial Intelligence*, 173(3-4) (2009): 466-500. For elaboration of the concept of autonomy to include adaptivity and precariousness, see Ezequiel Di Paolo, "Autopoiesis, Adaptivity, Teleology, Agency," *Phenomenology and the Cognitive Sciences,* 4 (2005): 97–125. For the enactive approach to participatory sense-making and social cognition, see Hanne De Jaegher and Ezequiel Di Paolo, "Participatory Sense-Making: An Enactive Approach to Social Cognition," *Phenomenology and the Cognitive Sciences*, 6(4) (2007): 485-507. For discussion of the relationship between the enactive approach and the extended mind or extended cognition theory, see Ezequiel Di Paolo, "Extended Life," *Topoi* 28 (2009) 9–21, and Evan Thompson and Mog Stapleton,

"Making Sense of Sense-Making: Reflections on Enactive and Extended Mind Theories," *Topoi* 28 (2009): 23-30.

Notes

[1] Dreyfus (1972) is probably the earliest statement of what is now known as embodied cognitive science. Yet this work was largely critical. Constructive theories and models of embodied cognition started to arrive in the mid-1980s, in works by Winograd and Flores (1986), Lakoff (1987), and Johnson (1987).

[2] Wilson and Foglia (2011) states the Embodiment Thesis as follows: "Many features of cognition are embodied in that they are deeply dependent upon characteristics of the physical body of an agent, such that the agent's beyond-the-brain body plays a significant causal role, or a physically constitutive role, in that agent's cognitive processing."

[3] Note that the human body as a whole comprises a number of overlapping autonomous systems, such as the nervous system and the immune system, and that these systems can incorporate extra-organic elements (e.g., neural prosthetics) into their operationally closed networks of enabling processes (Di Paolo 2009; Thompson and Stapleton 2009). For discussion of the relation between these systemic characterizations of the body and the phenomenological or subjective and experiential aspects of the body, see Thompson (2007).

References

Alexandre, G. and Zhulin, I. B. (2001) "More than One Way to Sense Chemicals*,"* *Journal of Bacteriology* 183: 4681-4686.

Barandiaran, X. E. and Egbert, M. D. (2013) "Norm-Establishing and Norm-Following in Autonomous Agency," *Artificial Life,* in press. doi:10.1162/ARTL_a_00094.

Clark, A. (2008) *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*, New York: Oxford University Press.

de Haan, S. and Fuchs, T. (2010) "The Ghost in the Machine: Disembodiment in Schizophrenia. Two Case Studies," *Psychopathology* 43: 327-333.

De Jaegher, H. and Di Paolo, E. (2007) "Participatory Sense-Making: An Enactive Approach to Social Cognition," *Phenomenology and the Cognitive Sciences*, 6(4), 485-507.

De Jaegher, H. and Di Paolo, E. (2013) "Enactivism is not Interactionism," *Frontiers in Human Neuroscience* 6(345). doi: 10.3389/fnhum.2012.00345.

De Jaegher, H., Di Paolo, E., and Gallagher, S. (2010) "Can Social Interaction Constitute Social Cognition?" *Trends in Cognitive Sciences* 14(10): 441-447. doi:10.1016/j.tics.2010.06.009.

Di Paolo, E. A. (2005) "Autopoiesis, Adaptivity, Teleology, Agency," *Phenomenology and the Cognitive Sciences* 4: 97–125.

Di Paolo, E. (2009) "Extended Life," *Topoi* 28: 9–21.

Di Paolo, E., and De Jaegher, H. (2012) "The Interactive Brain Hypothesis," *Frontiers in Human Neuroscience,* 6(163) doi:10.3389/fnhum.2012.00163

Dreyfus, H. (1972) *What Computers Can't Do*. New York: Harper & Row.

Egbert, M. D. (2013) "Bacterial Chemotaxis: Introverted or Extroverted? A Comparison of the Advantages and Disadvantages of Basic Forms of Metabolism-Based and Metabolism-Independent Behavior Using a Computational Model," *PLoS ONE* 8(5), e63617.

Egbert, M.D., Barandiaran, X. E., and Di Paolo, E. A. (2010) "A Minimal Model of Metabolism-Based Chemotaxis," *PLoS Computational Biology*, 6(12):e1001004.

Egbert M. D., Barandiaran X. E., Di Paolo, E. A. (2012) "Behavioral Metabolution: The Adaptive and Evolutionary Potential of Metabolism-Based Chemotaxis," *Artificial Life* 18(1): 1-25.

Froese, T. and Ziemke, T. (2009) "Enactive Artificial Intelligence: Investigating the Systemic Organization of Life and Mind," *Artificial Intelligence* 173(3-4): 466-500.

Fuchs, T. (2012) "Are Mental Illnesses Diseases of the Brain?" in S. Choudhury and J. Slaby (eds.) *Critical Neuroscience: A Handbook of the Social and Cultural Contexts of Neuroscience*, pp. 331-344. New York: Wiley-Blackwell.

Gallese, V. (2010) "Embodied Simulation and its Role in Intersubjectivity," in T. Fuchs, H. C. Sattel, and P. Henningsen (eds.), *The Embodied Self. Dimensions, Coherence and Disorders*, pp. 78-92. Stuttgart: Schattauer.

Goldman, A. (2012). A Moderate Approach to Embodied Cognitive Science," *Review of Philosophy and Psychology* 3 (1): 71-88.

Goldman, A. and de Vignemont, F. (2009) "Is Social cognition Embodied?" *Trends in Cognitive Sciences* 13(4): 154-9.

Hutto, D. and Myin, E. (2013) *Radicalizing Enactivism: Basic Minds without Content*, Cambridge, MA: MIT Press.

Johnson, M. (1987). *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason.* Chicago: University of Chicago Press.

Kyselo, M. and Di Paolo, E. (under review). "Locked-in Syndrome: A Challenge for Embodied Cognitive Science," *Phenomenology and the Cognitive Sciences*.

Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind.* Chicago: University of Chicago Press.

Lakoff, G. and Johnson, M. (1999) *Philosophy in the Flesh: the Embodied Mind and its Challenge to Western Thought*. New York: Basic Books.

Maturana, H., and Varela, F. (1980). *Autopoiesis and Cognition: The Realization of the Living*. Dordecht: D. Reidel Publishing Co.

Merleau-Ponty, M. (1963) *The Structure of Behavior*, trans. A. Fisher. Pittsburgh, PA: Dusequesne University Press.

Noë, A. (2004). *Action in Perception*. Cambridge, MA: MIT Press.

Núñez R. (2010) "Enacting Infinity: Bringing Transfinite Cardinals into Being," in J. Stewart, O. Gappene, and E. Di Paolo (eds.), *Enaction: Towards a New Paradigm in Cognitive Science,* pp. 307-333. Cambridge, MA: MIT Press.

O'Regan, K. and A. Noë. (2001) "A Sensorimotor Account of Vision and Visual Consciousness," *Behavioral and Brain Sciences* 24: 939–1031.

Shapiro, L. (2011) *Embodied Cognition*. New York: Routledge.

Thompson, E. (2001) "Empathy and Consciousness," *Journal of Consciousness Studies* 8, No. 5-7 (2001): 1-32.

Thompson, E. (2011a) "Précis of *Mind in Life*," *Journal of Consciousness Studies* 18 (2011): 10-22.

Thompson, E, (2011b) "Reply to Commentaries," *Journal of Consciousness Studies* 18 (2011): 176-223.

Thompson, E. (2011c) "Living Ways of Sense-Making," *Philosophy Today* SPEP Supplement 2011: 114-123.

Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: Harvard University Press.

Thompson, E. and Cosmelli, D. (2011) "Brain in a Vat or Body in a World? Brainbound versus Enactive Views of Experience," *Philosophical Topics* 39: 163-180.

Thompson, E. and Stapleton, M., (2009) "Making Sense of Sense-Making: Reflections on Enactive and Extended Mind Theories," *Topoi* 28 (2009): 23-30.

Thompson, E., and F. Varela (2001) "Radical Embodiment: Neural Dynamics and Consciousness," *Trends in Cognitive Sciences*, 5, 418–425.

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

Varela, F. J., Thompson, E., and Rosch, E. (1991). *The Embodied mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

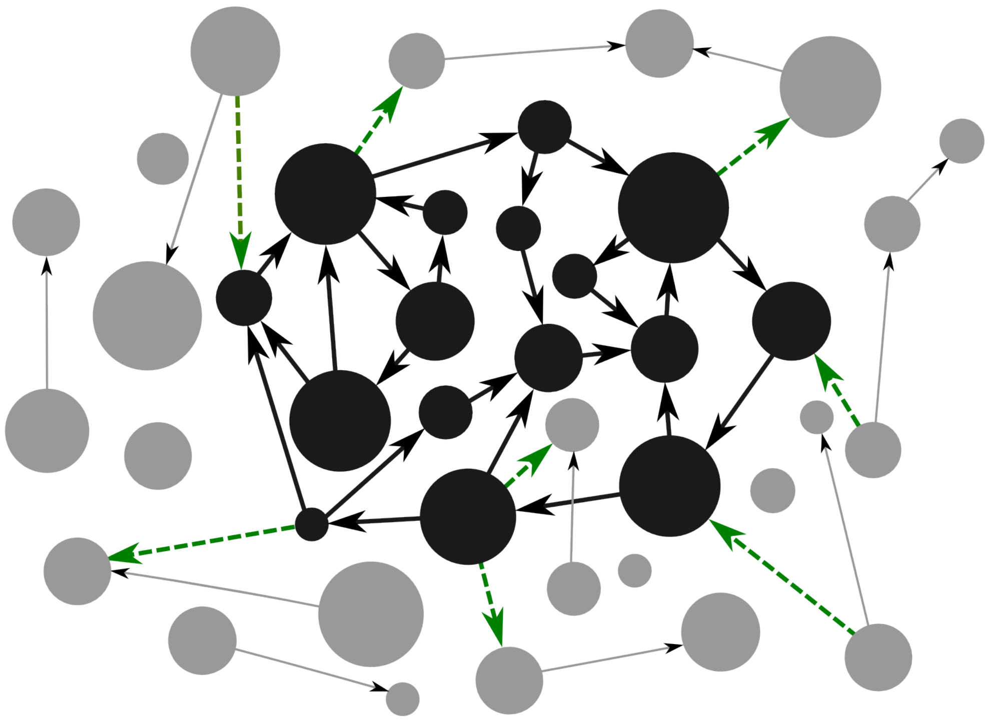
Weber, A. and Varela, F. J. (2002) "Life after Kant: Natural purposes and the Autopoietic Foundations of Biological Individuality," *Phenomenology and the Cognitive Sciences* 1: 97–125.

Wheeler, M. (2005) *Reconstructing the Cognitive World: The Next Step*. Cambridge, MA: MIT Press.

Wheeler, M. (2010) "In Defense of Extended Functionalism," in Richard Menary (ed), *The Extended Mind.* Cambridge, MA: MIT Press.

Wilson, R. A. and Foglia, L. (2011) "Embodied Cognition," *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2011/entries/embodied-cognition/>.

Winograd, T. and Flores, F., (1986) *Understanding Computers and Cognition: A New Foundation for Design,* Norwood, NJ: Ablex.



Copyright Ezequiel Di Paolo, 2013. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.
http://creativecommons.org/licenses/by-nc-sa/3.0/deed.en_US

Figure 1: A schematic illustration of the concept of operational closure. Circles are observed processes and arrows indicate enabling conditions between processes. The black circles form part of an operationally closed network of enabling relations. Each black circle has at least one arrow arriving at it and at least one arrow coming from it respectively originating or terminating in another black circle. Dashed arrows indicate enabling relations between processes in the operationally closed network and processes that do not belong to it.