

Living Systems: Autonomy, Autopoiesis and Enaction

Mario Villalobos · Dave Ward

Received: 1 July 2013 / Accepted: 22 January 2014 / Published online: 12 February 2014
© Springer Science+Business Media Dordrecht 2014

Abstract The autopoietic theory and the enactive approach are two theoretical streams that, in spite of their historical link and conceptual affinities, offer very different views on the nature of living beings. In this paper, we compare these views and evaluate, in an exploratory way, their respective degrees of internal coherence. Focusing the analyses on certain key notions such as autonomy and organizational closure, we argue that while the autopoietic theory manages to elaborate an internally consistent conception of living beings, the enactive approach presents an internal tension regarding its characterization of living beings as intentional systems directed at the environment.

Keywords Autonomy · Organizational closure · Autopoietic theory · Enactive approach

1 Introduction

We do not think of rivers as moving according to their own purposes, or lightning as choosing its path to the ground. Nor do we say that the flame of a candle consumes the wax because it seeks to preserve its existence, or that it values the presence of oxygen because it contributes to its existence. We can easily see that such anthropomorphic biases are, in general, inappropriate when describing or explaining natural phenomena. If the teleological language of purpose, agency and valuation is usually inappropriate in scientific descriptions and explanations of natural phenomena, should we use it when studying the subclass of natural phenomena made up of living beings? Is there something radically different about living beings that sets them apart from the rest of natural phenomena, and necessitates the use of teleological explanatory and descriptive tools?

M. Villalobos (✉) · D. Ward
School of Philosophy, Psychology and Language Sciences, University of Edinburgh,
Edinburgh EH8 9AD, UK
e-mail: mario.kirmayr@gmail.com

D. Ward
e-mail: Dave.Ward@ed.ac.uk

In this paper, we compare two closely related approaches to the study of life and cognition—the enactive approach¹ (Varela et al 1991; di Paolo 2005; Thompson 2007) and the autopoietic approach (Maturana 1975, 1981, 1987, 2003)—and consider the different answers that they give to these questions. Doing so allows us to identify a problem at the heart of the enactive approach. We will see that both the enactive and autopoietic approaches begin from the observation that living systems are in a particular kind of dynamic and reciprocal relationship with their environments. Comparing the different consequences that enactivists and autopoieticists draw from this observation, lets us see that the enactivist’s invocation of the kinds of teleological concepts described above is not legitimated by the appeal to the dynamics of organism/environment interactions in the way that enactivists have claimed.

Before beginning, some provisional sketches of the respective attitudes to the study of life and mind held by autopoietic and enactive approaches will help to set the scene. The autopoietic theory adopts what we take to be a strict naturalistic attitude, which says roughly: ‘Life is a natural phenomenon and must be studied as such, that is, by appealing to the same ontological assumptions and explanatory principles that we use to study any natural phenomenon in general. These ontological assumptions are, among others, that natural phenomena are what they are and occur as they occur (1) without having any goals, purposes or intentions; (2) without having any freedom to choose or select what to do and (3) without being governed by any normative consideration in terms of what may be good or bad, beneficial or harmful’.

The enactive approach, instead, adopts what we will call an existentialist attitude (given its inspiration in Jonas’ phenomenology of life, and through him, indirectly, in Heidegger’s existential phenomenology). The existentialist attitude says something like this: ‘Since we humans are living beings, we know life from within—from the perspective of our lived experience—which is an epistemic privilege that we must not dismiss. We experience ourselves as free agents capable of making choices according to our own norms and values, as controlling our behaviour according to our own goals and purposes, and as responding to a world that is not an indifferent environment, but a meaningful horizon of possibilities. We are, in sum, ‘existential beings’ (in the Heideggerian sense). Given that other organisms are living beings just as we are, why should we assume that they lack these properties that we experience as being involved in our living?’

This attitude, whose philosophical grounds we will discuss in the next section, assumes that all living beings, and only living beings, share a common form of existence that involves teleological properties and capacities such as purposefulness, intentionality, possibilities of action, agency, self-control, normativity, meaning, etc.

¹ The label ‘enactivism’ has been applied to a wide variety of views. Some views have been dubbed enactivist only in virtue of their emphasis on the coupled dynamics of organism/environment interactions in their accounts of cognition (including Hurley 1998; Noë 2004; and Hutto and Myin 2013). Our discussion of enactivism here applies not to those views, but only to enactivist views that, in addition, accord central importance to the autopoietic organization of living systems. The most notable exponents of this kind of view are Francisco Varela (see, e.g. Varela 1991, Varela and Weber 2001), Evan Thompson (see, e.g. Thompson 1995, Thompson 2007) and Ezequiel di Paolo (see, e.g. di Paolo 2005). All subsequent references to enactivism should be understood as pertaining only to this latter autopoiesis-emphasising group of theories. Other enactivist treatments that fall under the scope of our critique here includes Moreno and Barandarian (2004), Froese and Ziemke (2009) and the papers collected in Stewart et al. (2010).

and that justifies the use of special descriptive categories that cannot be applied to inorganic natural phenomena.

Faced with these alternatives, we might want to ask: ‘Which attitude, the naturalistic or the existentialist, is the correct one for the study of living beings?’, or ‘Which theory, the autopoietic or the enactive, tells us the truth about living beings?’ Though our argument in this paper certainly bears on these important questions, it is not part of our aim to settle them here. In part, this is because we think one can always conceive of a common and broader conceptual space in which both attitudes and theories may tell us, in their own terms, something sensible and useful about the phenomenon of life. Our purpose in this paper is instead to compare aspects of the conceptual foundations of the autopoietic theory and the enactive approach in a way that brings out a problem for the way in which the enactivist justifies their ‘existential’ attitude to the study of life and cognition, and to highlight that this problem does not occur for the autopoietic view. The autopoietic view may nonetheless be misguided, or the enactivist’s ‘existential’ attitude may nonetheless be legitimated on other grounds.

Before proceeding, a few words about the structure of the paper. In a standard (historical or genealogical) presentation of the two views and their relations, the autopoietic theory should come first and the enactive approach second. However, for our purposes, it is better to begin with a brief reconstruction of the foundations of the enactive approach (Section 2), focused on its key concept of *autonomy*, and the way in which the enactivist thinks that this property of living systems derives from the dynamics of their structural coupling with the environment. We then review, in Section 3, the autopoietic approach from which enactivists inherit their view of the dynamics of organism/environment interactions. The aim here is not to show that certain central concepts of the enactive approach have their roots in the autopoietic theory, which is a historical fact that, per se, offers no philosophical problem. Instead, presenting the two views in this way allows us to see clearly, in Section 4, that the enactivist’s existential attitude toward life and cognition does not follow from the view of the dynamics of organism/environment relations that they share with the autopoietic theorist.

2 The Enactive Notion of Autonomy

The notion of autonomy plays a central role in the enactive approach. ‘A distinctive feature of the enactive approach is the emphasis it gives to autonomy’, says Thompson in his *Mind in Life* (2007, p. 37), whilst Di Paolo, Rohde and De Jaegher present autonomy as one of the ‘five highly intertwined concepts that constitute the basic enactive approach’ (2010, p. 37). The enactive notion of autonomy comes originally from Varela’s work (1979) on biological autonomy and self-referential calculus (mathematical formalizations of the concept of organizational closure), but the particular significance this notion has for the enactivist is usually brought out (see Thompson 2007, chapters 3 and 4) by combining it with two complementary theoretical approaches as follows: (1) Dynamical Systems Theory (DST), and (2) Merleau-Ponty’s phenomenology of perception and action. As we shall see below, it is in the context of these approaches that the notion of autonomy takes on its distinctive importance for the enactive approach. For the enactivist, autonomy marks a fundamental and exclusive

property of living systems; their being centres of intentional directedness and normative evaluation. In what follows, we will analyse the notion of autonomy as presented in what we take to be the most comprehensive, canonical and updated exposition of enactivism thus far: Evan Thompson's *Mind in Life* (2007).

There, Thompson speaks of autonomy in two different but complementary senses as follows: (1) autonomy as self-determination, and (2) autonomy as organizational (or operational) closure. The first sense of autonomy is familiar.

An autonomous system is a self-determining system, as distinguished from a system determined from the outside, or a heteronomous system. On the one hand, a living cell, a multicellular animal, an ant colony, or a human being behaves as a coherent, self-determining unity in its interactions with the environment. An automatic bank machine, on the other hand, is determined and controlled from the outside, in the realm of human design (Thompson 2007, p. 37).

Autonomy is self-determination or self-government. A system is autonomous if it follows its own rules; if no external entity specifies, controls or commands its behaviour. A bank machine, and in general any man-made machine, operates following instructions given by an external agent, whereas living beings operate as self-determining systems that do not admit this kind of external command. This notion of autonomy has normative connotations (on which more below), as suggested by the word's original social-political root: '*Autonomy* and *heteronomy* literally mean, respectively, self-governed and other-governed' (Thompson 2007, p. 43 original emphasis).

The second sense of autonomy is more technical, and lacks a normative dimension. Here, autonomy refers to the way in which the constituent processes of a dynamic system are organized. 'In complex systems theory, the term *autonomous* refers to a generic type of organization (where the) relations that define the autonomous organization hold between processes (...) rather than static entities' (Thompson 2007, p. 44; original emphasis). An autonomous system, in this second sense, is basically a system whose processes are organized in a circular manner, like a closed circuit; hence the idea of *organizational closure*. 'Every autonomous system is organizationally closed', says Thompson quoting Varela (Varela 1979, p. 58; Thompson 2007, p. 48). The formal definition of autonomy as organizational closure includes three related aspects.

In an autonomous system, the constituent processes (1) recursively depend on each other for their generation and their realization as a network, (2) constitute the system as a unity in whatever domain they exist and (3) determine a domain of possible interactions with the environment (Varela, 1979, p. 55, quoted by Thompson 2007, p. 44).

As examples of this kind of system, enactivists point out single cells, multicellular organisms, nervous systems, immune systems, microbial colonies, ecosystems, etc. (in general, any system that is a living being or whose relevant components are living beings). All such closed systems, says Thompson (2007), now combining both senses of autonomy, 'need to be seen as sources of their own activity, specifying their own domain of interactions, not as transducers or functions for converting input instructions into output products' (p. 46).

2.1 Sensorimotor Loops: Without Starting and End Points

Interestingly, when Thompson opens the chapter about autonomy in his book *Mind in Life*, he does it by reviewing some central concepts of DST. According to his view, ‘some basic ideas about dynamic systems (...) form a background for the enactive approach’ (Thompson 2007, p. 37). This link, as we shall see, is not trivial. For Thompson, ‘the dynamicist idea that cognitive processes are always ongoing with no clear starting or end points’ is a key one (2007, p. 43), and will determine in part the manner in which he reads the notion of ‘closure’, The circular (recursive) organization of autonomous systems challenges the linear and sequential descriptions of cognitive processes.² In an ongoing circular network, starting and end points cannot be non-arbitrarily specified. Enactivists hold that the nervous system is an autonomous system, and hence an organizationally closed system. As a result, there is no non-arbitrary way of specifying starting and end points within its ongoing sensorimotor dynamic. To illustrate this idea, Thompson recurs to Merleau-Ponty and quotes a famous passage from *The Structure of Behaviour*.

It is clear that each of my movements responds to an external stimulation; but it is also clear that these stimulations could not be received without the movements by which I expose my receptors to their influence ... it is impossible to say “which started first” in the exchange of stimuli and responses. Since all movements of the organism are always conditioned by external influences, one can, if one wishes, readily treat behaviour as an effect of the milieu. But in the same way, since all the stimulations which the organism receives have in turn been possible only by its preceding movements which have culminated in exposing the receptor organ to the external influences, one could also say that the behaviour is the first cause of the stimulations (Merleau-Ponty 1963, p. 13, quoted by Thompson 2007, p. 48).

In the sensorimotor loop of the organism, perception and action are inseparable, in the sense that they are co-dependent aspects of a single circular process. As Merleau-Ponty notes, an observer can, according to her own descriptive purposes, stipulate starting and end points within this process. We can, for example, stipulate that the sensorimotor process begins with sensory stimulation and ends with motor response. There is nothing inherently wrong with such stipulations. However, attention to the circular organisation of the sensorimotor process shows that these stipulations do not reveal intrinsic properties of its dynamics. It would be equally valid to stipulate that the process begins with motor response and ends with sensory stimulation. Whether we see the dynamics of the system as running in one direction (sensation→action) or the other (action→sensation) has to do with our observational conventions, not with the organization of the system. That is why Merleau-Ponty says that in the exchange of stimuli

² This point is also of central importance to Hurley’s work on perception, agency and consciousness (Hurley 1998). Importantly, Hurley does not combine this point with the ‘existentialist’ attitude we briefly sketched above (and will explore in more detail below), so her view does not suffer from the tension that we argue arises for the autopoietic enactivist.

and responses, ‘it is impossible to say which started first’; we can certainly *assign* starting and end points, but we cannot *find* them.

2.2 From Sensorimotor Loops to Meaningful Worlds

As will become clear in the next section, there is no significant divergence between enactivism and its autopoietic ancestors as far as the above account of the dynamic circular organisation of cognitive systems is concerned. The divergence, which will ultimately result in a tension within the enactive approach that fails to arise for its autopoietic ancestor, comes when the enactivist claims that the circular structure of sensorimotor processes has specific phenomenological ramifications. After analysing the circularity of autonomous systems, Thompson adds an important claim about the consequences of the closed nature of their organization. Autonomous systems, thanks to their closure, emerge as ‘selves’, and in doing so, transform a (neutral) environment into a meaningful (enacted) ‘world’.

Whereas autopoietic closure brings forth a ‘bodily self’ at the level of cellular metabolism, sensorimotor closure produces a ‘sensorimotor self’ at the level of perception and action(...) In both cases, we see the co-emergence of inside and outside, of selfhood and correlative world or environment of otherness, through the generic mechanism of network closure (autonomy) (Thompson 2007, pp. 48–49).

This idea is developed in several ways in Thompson’s treatment, but the key claim is that living beings constitute a ‘perspective’ or ‘point of view’ according to which the (merely physical) surrounding emerges as a significant world, as an *Umwelt* (the term of von Uexküll, a favourite author in the enactive literature). According to the enactivist, a living being is a biological self that specifies its own identity as something different from, though correlative to, the outside world. And the enactivist claim is that it can do this in virtue of the circular organisation of its sensorimotor dynamics, as described above.

In the case of animal life, the environment emerges as a sensorimotor world through the actualization of the organism as a sensorimotor being. The organism is a sensorimotor being thanks to its nervous system. The nervous system connects anatomically distant sensory and motor processes, subsuming them in operationally closed sensorimotor networks. Through their coherent, large-scale patterns of activity, these networks establish a sensorimotor identity for the animal—a sensorimotor self. In the same stroke, they specify what counts as ‘other’, namely, the animal’s sensorimotor world (Thompson 2007, p. 59).

At this point, we are already in phenomenological terrain. The co-emergence of the biological self and its correlative world yields an intentional relation between the organism and its environment. In keeping with the phenomenological tradition, this intentional relation is understood by the enactivist not in terms of, ‘*states having content*’ but (in terms of) ‘*acts having directedness*’ (Thompson 2007, p. 25; original emphasis). The organism embodies an intentional experience that, strictly speaking, has no content (in the representational sense) but rather a particular orientation towards its

world. Rather than *representing* an independent world, the organism is *oriented towards* a correlative and co-emergent world. This phenomenological step is crucial. Since the organism has a ‘perspective’ on the environment, the environment appears no longer as a neutral and indifferent field of potential action, but instead as a field of relevance wherein certain things are important, desirable or good for the organism and others not.

How, according to the enactivist, do a meaningful environment and an intentionally-directed organism arise from the circular dynamics of sensorimotor loops? The key lies in the ways in which the particular embodiment of the organism, and of the nervous system that modulates its sensorimotor coupling with the environment, places constraints both on the portions of the world with which the organism can couple, and on the way in which those selective couplings will unfold. Thompson (following Varela 1991) walks us through a simple case.

‘Consider motile bacteria swimming uphill in a food gradient of sugar. These cells tumble about until they hit on an orientation that increases their exposure to sugar, at which point they swim forward, up-gradient, toward the zone of greatest sugar concentration... These bacteria are autopoietic and embody a dynamic sensorimotor loop: the way they move (tumbling or swimming forward) depends on what they sense, and what they sense depends on how they move. This sensorimotor loop both expresses and is subordinated to the cell’s autonomy... As a result, every sensorimotor interaction and every discriminable feature of the environment embodies or reflects the bacterial perspective. Thus, although sucrose is a real and present condition of the physicochemical environment, its status as food is not. That sucrose is a nutrient is not intrinsic to the sucrose molecule, but is a relational feature, linked to the bacterium’s metabolism. Sucrose has significance or value as food, but only in the milieu that the organism itself enacts. Thus, thanks to the organism’s autonomy, its niche has a “surplus of significance” compared with the physicochemical environment’ (Thompson 2005, p.418).

In virtue of the structural properties of the bacterium and the part of the physical world in which it is located, the sensorimotor dynamics of the bacterium/environment coupling implicate only a subsection of the states and properties of the world. Thus, for the enactivist, the circular nature of the organism’s sensorimotor coupling with the environment grounds the co-emergence of properties in the environment that are meaningful to the organism (such as the sucrose’s property of ‘being a nutrient’) and properties of the organism involving directedness toward its environment (such as its *needing* or *aiming towards* sucrose). This is what the enactivist has in mind when she claims that the autonomous character of an organism’s relation to its environment allows it to enact, or to bring forth, a meaningful world. The idiosyncracies of the organism’s embodiment lead to dynamic coupling with some aspects of the world rather than others, and it is in virtue of this fact that the world can appear to the organism as a value-laden environment toward which it is directed.

It is this commitment of enactivism that lies behind its claim to have uncovered the origins of our mind in basic biological processes, and its commitment to the strong continuity of life and mind. The commitment is picked up on and developed in different

ways by different enactivists; for example, via the concept of sense-making (Weber and Varela 2002), or of adaptivity (Di Paolo 2005). We will not review these concepts here, though. It is enough, for the purposes of our discussion, that we have followed the reasoning through which the enactive approach, starting with the definition of living beings as autonomous systems, arrives at the view that living beings constitute identities or selves that specify their own outside (but not independent) world, toward which they are directed in a normative way.

3 Autopoiesis: Autonomy?

In the next section, we will explain why we think that the role played by the concept of autonomy for the enactivist results in a problematic tension for the view. But to do so, it will help to contrast the conclusions that the enactivist draws from a system's possession of circular structural dynamics with those drawn by the autopoietic theory. In the autopoietic theory (Maturana 1975, 1981, 1987, 2003), autonomy is a concept that appears occasionally in the description of living beings, but that is neither accorded a privileged status nor invoked as a part of the theory's 'conceptual core'. Nonetheless, the concepts of self-determination and organizational closure that, as we have just seen, underpin the enactive notion of autonomy do appear in the autopoietic literature and play important roles. In contrast to the enactivist, however, Maturana does not unify the notions of self-determination and organizational closure under the concept of autonomy (or under any concept at all) because, as we shall see, in the autopoietic theory these notions are entirely independent.

What enactivists call self-determination, i.e. the fact that the system does not admit external specifications or determinations, corresponds basically to what Maturana (1975, and subsequently) calls 'structural determinism'. A structurally determined system is a system within which everything that happens is determined or specified by its own structure—by its own constitution or composition. External factors or interactions can trigger or initiate certain structural changes in the system (reactions, processes, behaviours) but cannot specify or instruct them. The ways in which structurally determined systems change, react and evolve over time are determined by their particular structural dynamics, which cannot be altered by things external to the system.³ Now, under this definition, which systems are structurally determined and which are not? Here is Maturana unpacking the idea.

The structural changes that a composite unity [in Maturana's jargon 'composite unity' means 'system'] undergoes as a result of an interaction are [...] determined by the structure of the composite unity. Therefore, an external agent which interacts with a composite unity only triggers, rather than determining in it a structural change. Since this is a constitutive condition for composite unities,

³ Note that this does not involve the obviously false claim that structurally determined systems cannot undergo change as a result of perturbations from their environment—the claim is rather that such changes will always unfold according to the fixed structural dynamics of the system.

nothing external to them can specify what happens in them: there are no instructive interactions for composite unities (Maturana 2003, p. 61).

Here, Maturana speaks of ‘composite unities’ in general, without making any distinction within them. This is because, according to Maturana, every physical system, whether static or dynamic, alive or not, natural or artificial, is a structurally determined system (Maturana 1975, 1981, 1987, 2003). For the autopoietic theory, then, systems are autonomous with respect to the specification of their changes of state *regardless* of whether their dynamics have the closed structure identified in the previous section. To say that living beings are autonomous systems, in the sense that the ways in which they change in response to external perturbations are not determined or governed by external factors, is, from this point of view, ontologically trivial and marks no interesting difference with non-living systems.

The distinction that Thompson makes between living systems and machines (man-made artefacts), i.e. between autonomous and heteronomous systems, thus has no place in the autopoietic theory. Indeed, when Maturana explains the notion of structural determinism by appealing to our quotidian experience, he uses man-made machines as his examples.

We know this [that systems are structurally determined] very well from using any of those pushbutton machines, in which if you push a button something happens—it washes, it glows, it plays music—which is not determined by your pushing the button, but, rather, is triggered by the pushing of the button [...] You do not instruct a system, you do not specify what has to happen in the system. If you start a tape recorder, you do not instruct it. You trigger it (Maturana 1987, p. 73).

So here, we have the first difference between the enactive approach and the autopoietic theory; while for the enactive approach, autonomy (self-determination) is a distinctive characteristic of *living* systems, possessed in virtue of their organizational closure; for the autopoietic theory, autonomy (structural determinism) is a constitutive property of *any* system, whether or not it is living or organizationally closed. However, the key difference between the enactive approach and the autopoietic theory, at least for our purposes in this paper, lies in the idea of autonomy as organizational closure.

3.1 Organizational Closure

In Section 2, we said that the enactivist’s notion of autonomy comes originally from Varela’s work on organizational closure in the 1970s. But the whole story is that Varela’s notion of organizational closure comes in turn from Maturana’s work on functional and organizational closure in the second half of the 1960s. The interesting point, as already noted, is not to establish a hierarchy of antiquity or originality, but to analyse the reach of the notion of organizational closure in the autopoietic theory and compare it with the enactive version.

It is (more or less) known that in the 1960s, before having the word ‘autopoiesis’ at hand, Maturana already conceived of living beings as circular (closed) systems of production—as self-producing systems. What is not so well known is that he conceived

of living beings as closed systems not only from the point of view of their productive chemical processes, but also, and more crucially, from the point of view of their *functional* organization as *sensio-effector* systems.⁴

The sensory and the effector surfaces that an observer can describe in an actual organism, do not make the [...] system an open [...] network because the environment (in which the observer stands) acts only as an intervening element through which the effector and sensory [surfaces] interact completing the closure of the system (Maturana 1975, p. 318).

The autopoietic claim that living systems are closed systems needs careful handling. The functional organization of the sensio-effector system of the living being is closed, but this is not because the environment is left outside the system. Rather, it is because the environment is always incorporated as a *functional step* within the system. The sensio-effector system is not closed *to* the environment; it closes on itself *through* the environment. Thus, the autopoietic view does not aim to deny the fact that, from a material and energetic point of view, living beings are essentially open systems (dissipative chemical structures) that interact with their environments—consuming nutrients and eliminating waste, for example. The sense of closure at issue for the autopoietic theorist is just the closure of sensorimotor loops through the environment that we saw the enactivist endorse in 2.1, above. The activities at the sensory and effector surfaces of an organism are co-dependent in the sense that the quoted passage from Merleau-Ponty indicates (p.8, above). What happens at the effector surfaces of an organism conditions, through the structure of the environment, what happens at its sensory surfaces. But what happens at its effector surfaces is likewise conditioned, through the structure of the organism, by what happens at its sensory surfaces. So what is closed here is the circuit of the organism's sensorimotor activity, a circuit which incorporates causal interchange with the environment.

Recall from Section 2.2 that, for the enactivist, it is just this circular nature of an organism's sensorimotor dynamics that results in the emergence of a meaningful environment for the organism, to which the organism is directed. The autopoietic theorist, however, draws a very different lesson from the closed dynamics of sensio-effector systems. Just as Merleau-Ponty noted that there is no non-arbitrary way of designating what comes first in the dynamic cycle of perception and action, the autopoietic theorist notes that, *as far as the closed structural dynamics of the system are concerned*, there is no non-arbitrary way of carving the sensorimotor loop up into distinct portions. In other words, what *we* see as the salient boundary between the internal states of the organism and the external states of its environment is something that, as far as the sensorimotor dynamics of the system are concerned, proves to be operationally 'transparent' and does not exist as such.

An observer that sees an effector/sensor correlation as an adequate behaviour does so because he or she beholds the organism in the domain of structural coupling in which the distinguished behaviour takes place in the flow of its

⁴ See, e.g. Maturana's contribution to Maturana and Varela 1980, or his lovely autobiographical account in Maturana 2013.

conservation of adaptation. The organism in its operation [‘nonetheless’] does not act upon an environment; the environment exists only for an observer (Maturana 2003, pp. 102–103).

The autopoietic theorist thus holds that the ‘environment’ portion of the closed loop is something that exists only for an observer, rather than reflecting a boundary specified by the causal dynamics of the loop itself. The same applies to what we, as observers, can identify as the ‘organism’ portion of the loop. Maturana concludes that: ‘For a living system in its operation as a closed system, there is no inside or outside, (simply) it has no way of making the distinction’ (Maturana 2003, p. 99).

At this point, if the reader recalls the enactive version of the notion of autonomy, we can already see not only an important difference regarding the rendition of the concept of organizational closure, but a completely different conception of living beings and their ‘relationship’ with the environment. Let us recapitulate what we have seen until now.

So far, we have seen that both theories, autopoietic and enactive, subscribe to the view that living beings exhibit organizational closure. Nonetheless, starting from this point in common, these theories seem to arrive at very different conclusions about living beings. In the enactive case, we have the view of a system that, being autonomous, emerges as a self that is intentionally directed at a world, that defines an inside and an outside, and to which the environment appears endowed with significance and value. In the other case, we have the view of a system that does not distinguish itself from anything, to which there is no inside and outside, and to which something like ‘the environment’ or ‘the world’ does not exist at all.

4 No Direction, No Beginning, No End

So, the autopoietic and enactive theories begin from a shared emphasis on organizational closure, but arrive at opposing conclusions about directionality and teleology in living systems. Who should we side with? Our contention is that we should side with the autopoietic theorist here—the ‘existential’ attitude to living systems that we sketched in the introduction cannot be legitimated by appeal to organisational closure. The reasons for this are ones we have just reviewed on behalf of the autopoietic theorist—so let us consolidate them.

If we draw an arrow and ask you where it begins, where it ends and where it points, you should be able to answer easily. If we draw a circle and ask you the same questions, you should struggle. The simple explanation for this is that arrows have starting points, end points and directionality, whereas circles do not. The moral of the considerations about organizational closure in Sections 2.1 and 3.1, above, is that the structural dynamic of a living system is like a circle, rather than like an arrow. So if living systems are to be construed as being directed toward their environments, or if their cognitive processes are to be construed as having beginnings and ends, then this must be in virtue of something other than their organizational closure.

As we have seen Merleau-Ponty, Maturana, and Thompson agree, the dynamics of a continuous and circular process of sensorimotor correlations do not, in and of themselves, permit the non-arbitrary definition of starting and end points of that

sensorimotor process. We also saw above that an observer might, for the sake of certain descriptive purposes, *treat* sensorial stimulation as a starting point and the motor response as an end point. However, doing so would be a reflection of the interests of the observer, rather than of the dynamics of the system. If sensory or motor processes are to be construed as having beginnings and ends, then this must be in virtue of something other than their structural dynamics.

Relatedly, if an organism's sensorimotor system is closed in its dynamics in the way we have been discussing, then its sensory and motor surfaces are not entries and exits of the system, i.e. doors or windows through which the organism, considered as a sensorimotor entity, 'contacts' the environment. They are rather functional points that, in their interaction, close the organism as sensorimotor system *through* the environment, as discussed in Section 3.1. In terms of its functional organization (not in terms of its physical boundaries, of course), the organism has the environment as one more of its functional components; not as something external to it. In other words, what we see as the distinction between the environment and the organism is not a distinction we find within the organism's closed sensorimotor dynamics. Now, if the functional organization of the organism's sensorimotor interactions with its environment is strictly circular in this way, then each portion of the sensorimotor loop that we might single out fulfils exactly the same role—as a link in the larger loop through which the system closes on itself. Can we identify, within such a system, a particular part of the sensorimotor loop toward which the rest of the system is *directed*? If the portion of the loop that we, from our position as observers, see as the environment of the organism is only one seamless link in the circular dynamics of the system, then what is the ground to say that this portion, rather than any other, is *the* one around which the rest of the system works, *the* part to which the rest of the system points? When we say that the sensorimotor organization of the organism is circular, what we say is that said network is, *as far as its structural dynamics are concerned*, strictly symmetric and equivalent in all its portions. If we restrict our focus to an organism's sensorimotor dynamics, then there is no non-arbitrary reason to single out one portion of its looping structure to which it is directed. In other words, if the sensorimotor activity of an organism is to be construed as directed toward its environment, then this must be in virtue of something other than its structural dynamics.

To emphasise once more, as observers, we can and do pick out particular portions of the sensorimotor loops of living systems as points of reference and describe systems as oriented toward them. Indeed, since we as observers are actually located in one portion of the sensorimotor network, namely the gap that we see between the sensory and motor surfaces of the organism, we, tacitly (as a sort of 'natural attitude'), always choose that node as a point of reference. We see the organism's functioning as directed precisely at the region in which we operate as observers, as oriented to what we see existing between the sensory and motor surfaces of the organism; that is, we naturally see an organism's functioning as directed toward the environment in which we observe it. Such a perspective is only objectionable to an autopoietic theorist if we also assume that the directionality we see is a property of the organism's own sensorimotor activity. Considered in itself, that sensorimotor activity is organizationally closed, and so without direction. We can compare the question of the direction or orientation of a closed sensorimotor system to the question of whether a moving object is approaching or moving away. What is the answer that best reflects the structural dynamics of the object itself? Neither answer will do this—seeing an object as moving away or as approaching

depends on the particular point of reference adopted by the observer. Just as an object's 'approaching' or 'moving away' are not properties that it can possess solely in virtue of its structural dynamics, an organism's orientation to its world (if it has one) cannot be derived from the properties of its operationally closed sensorimotor interactions.

The reader might have noticed that we have been conspicuously silent about the feature of the enactivist view that seems most important here—we saw in Section 2.2 that, for the enactivist, an organism's orientation toward a meaningful environment was possessed in virtue of the way the physical realization (or embodiment) of the organism and of its environment constrain its sensorimotor activity. The sucrose in the bacterium's environment becomes a *nutrient*, and something which the bacterium can *need*, or *aim towards*, in virtue of the structural properties of sucrose and the bacterium's embodiment, and the ways that those conspire to shape the bacterium's sensorimotor coupling with its environment. Though this is indeed the key step in the enactivist's legitimation of their existential attitude toward living systems, we have postponed our treatment of it until now because we think that appreciating the force of the points made above shows how it should be addressed.

Remember the list of examples with which we opened this essay—we do not think of rivers as moving according to their own purposes, or lightning as choosing its path to the ground or of candles as seeking out oxygen to consume to keep their flame alight. Why should the ways in which we think of living systems be any different? We saw in Section 2 that the enactivist answer appeals to the circular dynamics of the ways in which living systems selectively couple with aspects of their environment. But note that it is not only living systems that couple with some aspects of their environment rather than others. A burning candle remains alight as long as the physical materials needed for the process of combustion are available in the environment. The continued existence of the candle's flame is determined by the dynamics of its coupling with some elements of the air that is its environment rather than others—oxidizer gases such as oxygen or nitrogen dioxide, and flammable gases such as ammonia and carbon monoxide. Inert gases such as nitrogen, carbon dioxide and argon are irrelevant to the preservation of the flame, and are not involved in its burning. The structural properties of the candle and its environment thus determine the ways in which it couples with some elements of its environment rather than others in ways that keep its flame alight. Nonetheless, this does not provide us with reason to say that the candle seeks out oxygen, or that oxygen appears to it as valuable. The candle's selective coupling with the environment does not suffice to endow it with a perspective, or to place it on a continuum with our minds and their orientation to the world, in the ways that the enactivist claims for the bacterium's coupling. This is not an isolated case. The water in Mario's mug absorbs material from the tea leaves but not the tea bag. The fabric of Dave's chair is shaped by the distribution of his mass on it, but not by the colour of his hair. Selective coupling does not suffice for a system to be directed toward its environment.

Perhaps it could be objected that the structural dynamics of these couplings do not have the appropriate circular organization required by the enactivist. At least in the case of the candle, we do not think this objection can be made. A burning candle is a circular chain of physicochemical events that maintains itself: the heat of the flame melts and liquefies the wax, the wax ascends through the wick (by capillarity) where it is vaporized to finally burn in the flame, whose heat melts and liquefies the wax, and so on. Here, we have a case of both organizational closure and selective coupling where a teleological description looks inappropriate. How about the tea leaves and the fabric of the chair? Composition of

Mario's tea leaves is certainly altered by the water as they in turn alter the water's composition, and the shape of the chair's fabric effects the way Dave distributes his mass, as well as his mass reciprocally effecting its shape. But perhaps here, we have only reciprocal causal interaction rather than genuine organizational closure. Here, some further explanation of how cases of organizational closure can be distinguished from cases of mere reciprocal causal interaction is required. But even if it could be shown that these systems lack appropriate organizational closure, we are owed an explanation of why this should matter. We saw earlier in this section that such organization does not suffice for a system's directedness to its environment. And we saw in the present discussion of the candle's flame that selective coupling does not suffice either. Perhaps they are, as the enactivist claims, jointly sufficient. But since there is good reason to believe that neither condition on its own implicates teleology, the enactivist needs to provide further explanation of why they should do so when combined.

5 Conclusion

The deep continuity of life and mind is a foundational (and, to many outside observers, counterintuitive) aspect of the enactivist view—so why has it seemed to enactivists that this continuity follows so straightforwardly from the properties of the structural dynamics of living systems? We conclude with the speculative suggestion that it is for reasons we have already mentioned above. As observers, it is clear to us that we can cleanly parse the dynamical loop of an organism/environment coupling into constituents of organism and environment. After all, the distinction between what is the organism, and what is its environment, is usually hard for us to miss. And once we have made this parsing, it is tempting to view the activity in the organismic portion of the loop as directed toward its environment, and the activity in the environmental part of the loop as directed toward the organism. It seems outlandish to deny, at least in the case of many complex organisms, that they do act in ways directed toward the world, and perceive in ways sensitive to their world. If we want to explain these facts, and are faced with the fact that the dynamics of a living system's sensorimotor couplings are organizationally closed, it is easy to map the distinctions between organism and environment, and between states of the organism directed toward the environment and states of the organism inherited from the environment, onto appropriate portions of those dynamics. But what we have argued in the foregoing is that undeniable as these distinctions may be to us as observers, they are not straightforwardly derived from the sensorimotor dynamics of living systems. As we noted at the outset, the enactivists' existential attitude toward living systems may yet be justified. But enactivists are wrong to think that this attitude, and the commitment to a deep continuity of life and mind that it entails, are justified by appeal to selective, organizationally closed, sensorimotor couplings with the environment. As we also noted at the outset, it was not our aim here to demonstrate the truth of the autopoietic approach to the study of living systems. However, we think that we have shown in the foregoing that as matters stand, with respect to their views of the ramifications of the structural dynamics of living systems, the autopoietic approach is to be commended over enactivism.

Acknowledgements We would like to thank three anonymous referees for their constructive comments and observations. Mario would like to thank Katja Abramova for helpful comments on an earlier draft of this paper.

References

- Di Paolo, E. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 4(4), 429–452.
- Di Paolo, E. A., Rohde, M., & De Jaegher, H. (2010). Horizons for the enactive mind: values, social interaction and play. In J. Stewart, O. Gapenne, & E. A. Di Paolo (Eds.), *Enaction: Towards a New Paradigm for Cognitive Science* (pp. 33–87). Cambridge, MA: MIT Press.
- Froese, T., & Ziemke, T. (2009). Enactive artificial intelligence: investigating the systemic organization of life and mind. *Artificial Intelligence*, 173, 466–500.
- Hurley, S. (1998). *Consciousness in action*. Cambridge, MA: Harvard University Press.
- Hutto, D., & Myin, E. (2013). *Radicalizing enactivism: basic minds without content*. Cambridge, MA: MIT Press.
- Maturana, H. (1975). The organization of the living: a theory of the living organization. *International Journal of Man-machine studies*, 7, 313–332.
- Maturana, H. (1981). Autopoiesis. In M. Zeleny (Ed.), *Autopoiesis: a theory of living organization* (pp. 21–33). New York: Oxford: North Holland.
- Maturana, H. (1987). Everything is said by an observer. In W. I. Thompson (Ed.), *GAI: A way of knowing* (pp. 65–82). Hudson, N.Y.: Lindisfarne Press.
- Maturana, H. (2003). The biological foundations of self-consciousness and the physical domain of existence. In N. Luhmann, H. Maturana, M. Namiki, V. Redder, & F. Varela (Eds.), *Beobachter: Convergenz der Erkenntnistheorien?* (pp. 47–117). München: Wilhelm Fink Verlag.
- Maturana, H. (2013). Reflections on my collaboration with Francisco Varela. *Constructivist Foundations*, 7(3), 155–164.
- Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and cognition: the realization of the living*. Dordrecht, Holland: Kluwer Academic Publishers.
- Moreno, A., & Barandarian, X. (2004). A naturalized account of the inside-outside dichotomy. *Philosophica*, 73, 11.26.
- Noë, A. (2004). *Action in perception*. Cambridge MA: MIT Press.
- Stewart, J., Gapenne, O., & di Paolo, E. A. (Eds.). (2010). *Enaction: towards a new paradigm for cognitive science*. Cambridge MA: MIT Press.
- Thompson, E. (2005). Sensorimotor subjectivity and the enactive approach to experience. *Phenomenology and the Cognitive Sciences*, 4, 407–427.
- Thompson, E. (2007). *Mind in life: biology, phenomenology and the sciences of mind*. Cambridge, MA: Harvard University Press.
- Varela, F. (1979). *Principles of Biological Autonomy*. New York: Elsevier North Holland.
- Varela, F. (1991). Organism: a meshwork of selfless selves. In A. Tauber (Ed.), *Organism and the origin of self* (pp. 79–107). Kluwer Academic Publishers: Dordrecht.
- Weber, A., & Varela, F. (2002). Life after Kant: natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, 1, 97–125.