

# Autopoiesis, adaptivity, teleology, agency

BY EZEQUIEL A. DI PAOLO

*COGS, Department of Informatics*

*University of Sussex, Brighton BN1 9QH, U.K.*

*Email: ezequiel@sussex.ac.uk, Fax: +44-1273-877873*

A proposal for the biological grounding of intrinsic teleology and sense-making through the theory of autopoiesis is critically evaluated. Autopoiesis provides a systemic language for speaking about intrinsic teleology but its original formulation needs to be elaborated further in order to explain sense-making. This is done by introducing adaptivity, a many-layered property that allows organisms to regulate themselves with respect to their conditions of viability. Adaptivity leads to more articulated concepts of behaviour, agency, sense-construction, health, and temporality than those given so far by autopoiesis and enaction. These and other implications for understanding the organismic generation of values are explored.

## 1. Introduction

Few philosophers have argued more passionately for the continuity between life and mind than Hans Jonas. His philosophy of the organism (Jonas, 1966, 1968, 1979) is extraordinarily rich in scope and boldly direct in its engagement with the central problems: what dynamical processes produce the special kind of identity of living organisms and how those same processes underlie landmark properties of mind such as inwardness and a concerned perspective on the world. Jonas has described the shape that a naturalization of teleology might have, both at its foundational grounding in metabolism as well as in the subsequent development of the polarities thereby established for the first time.

Yet his arguments and insights do not constitute a scientific theory – nor are they meant to. If they were to realize this potential it would be necessary to flesh them out in terms that are amenable to scientific formulation. There is no lack of suitable candidate frameworks, even if they do not figure prominently in mainstream contemporary biology. They form a set of mostly compatible but not fully unified ideas originating in the work of existentially inclined scientists such as Helmut Plessner, Frederik J. J. Buytendijk, Kurt Goldstein, Michael Polanyi, Paul Weiss, Howard Pattee, and Erwin Straus. The most recent and explicit proposal has been offered in an article by Andreas Weber and Francisco Varela (2002) where a link is offered between the theory of autopoiesis, a theory of minimal living organization proposed by Maturana and Varela in the 1970s, and sense-making, the instauration of a natural perspective from which encounters in the world are intrinsically meaningful for the organism following the norm established by the continuing process of self-production.

In this article, I shall argue that this proposal is a promising starting point but has a series of shortcomings that can be identified through scientific and phenomenological routes. If Varela and others have managed to see in autopoiesis a natural link with intrinsic teleology and sense-making, it is because they have complemented the theory with an additional assumption that has guided its interpretation. I will propose to make explicit this aspect of living beings: that of *adaptivity* or the capacity of an organism to regulate itself with respect to the boundaries of its own viability. A careful analysis of sense-making shows that different properties of adaptivity (self-monitoring, control of internal regulation, and control of external exchanges) are implied by assuming that organisms have a meaningful perspective on their world, hence this property is not simply

an *addendum* but is essential (together with autopoiesis) to naturalize sense-making. In turn, examining some of the implications of adaptivity sheds new light on different aspects of sense-making such as graded norms, indirect meaning construction, behaviour, agency, mal-adaptation, and temporality<sup>†</sup>.

## 2. Hans Jonas and the need for circulation

Weber and Varela (2002) argue that if we follow Kant’s exposition of natural teleology in his *Critique of Judgment*, there is room for doubt regarding its ontological implications. We see organisms as intrinsically teleological, i.e., made of components which are both the generative precursors and the end products of other such components, even if a causal mode of reasoning would be unable to lead us to this conclusion from first principles. But is this teleology real? Extending efficient causality to include the modern concept of self-organization does not dispel the possibility that the use of teleology is merely ascriptional. For Jonas, these are doubts worthy of a disembodied God-mathematician capable of indifferently switching between convenient descriptions that subsume a complex causal story (waves) and the causal story itself (oscillating particles). But our position as embodied observers is one of privileged standing with regard to the issue of teleology. We are physical beings *and* we have direct experience of our own inwardness and concern in our dealings with the world. This existential overturning of the question does not provide a solution to how teleology and causality are reconciled, but at least it forces us not to avoid the problem by ignoring the evidence of embodied experience. As Jonas puts it:

“How this finalism tallies, in the same world, with mechanical causality whose reality cannot be denied either is a problem not to be “solved” by sacrificing an evidence (purposiveness) to a theorem (exclusiveness of *causa efficiens*) which was derived by generalization from another evidence; but, if solvable at all, only by treating it as the profoundly challenging and as yet completely unsettled problem it is”, (Jonas, 1966, pp. 90-91).

Jonas shows his pragmatic inclinations by avoiding spending too much time on the possible but “useless” extremes of species-solipsism and panpsychism. He does so by means of arguments of parsimony and scientific plausibility. His essay on the philosophical aspects of Darwinism shows that the continuity of life linking humans to the rest of the natural world overcomes the Cartesian divide between automated animals and soulful persons, but the triumph of materialism thus achieved “contains the germ of its own overcoming” (Jonas, 1966, p. 53) for our own embodied experience does not withdraw, but trickles onto the natural world across the bridges provided by Darwin.

If we accept as plausible that the experience of concern is not exclusively human (though it may have some exclusive characteristics in humans) and that all other physical living beings may also be, rather than appear, intrinsically teleological, is this because they are living or simply because they are physical? For Jonas, again, it is a question of selecting the most informative option, the one that is more revealing in terms of the homology that can then be established with the structure of our experience. For

<sup>†</sup> I should make clear at this point that throughout this paper the term autopoiesis will be used in its original formulation and interpretation which are given in the primary literature and described in the next section. I recognize that there have been attempts to re-formulate this concept based on different kinds of worries. For Varela himself, although he always referred to the original definition, the concept has had a subtle but clear evolution noticeable in his later writings where the interpretation has increasingly emphasized the phenomenon of self-constituted and multi-layered processes of biological identity. Following the spirit of Weber and Varela’s conclusions, this paper can be partly read as an attempt to articulate what this later interpretation of autopoiesis implies in systemic terms intended to be compatible with those used in the original definition. If adaptivity is contained in new versions of autopoiesis (which is not immediately obvious), it still needs to be made explicit in operational terms.

Leibniz's monadology and for Whitehead's philosophy of the organism, Jonas argues (1966, pp. 95-96), (1968, pp. 235, 241), there is no *useful* concept of challenge to organic identity since this kind of identity is extended to cover all cases of physical identity, even that of particles that simply endure. Yet it seems a pragmatically vacuous extension of vocabulary to say that atoms die or molecules get sick. Such precariousness is given to organisms by their singular mode of identity: that of a stable dynamic form made of an ever-changing material substrate.

The break with the substantial mode of identity is achieved in *metabolism*, a self-affirming process of constant regeneration of form amidst a flux of matter and energy. This level of physical organization seems to have the necessary existential credentials: 1. the establishment of a distinct "self" for which being is its own doing and with physical and organizational distinctions between inside and outside, 2. an entity which is in constant environmental challenge, in need of material turnover and with the freedom to achieve it, and 3. the establishment of a normativity following the logic of metabolism according to which otherwise neutral events, both internal and external, can be good or bad for the continuation of the organism. Jonas's proposal is that metabolism *is* intrinsically teleological, a statement that cannot be arrived at by the unprepared, disembodied observer.

From the viewpoint of science, if we were to take Jonas seriously and use his insights to guide further research (for instance, in the synthesis of intrinsically teleological artefacts), some problems of method present themselves. While we cannot doubt the fact that we care about our own existence, the extension of teleology to metabolism is justified partly by means of intuition outside scientific discourse. Jonas implicitly admits that establishing in metabolism the breaking point between extended neutral processes and concerned identity, is a matter of appropriate choice. How are we to justify this and further choices, or question their sufficiency, when the criteria of validation are, at least partly, outside science? The answer must be: by the use of phenomenological insight or other disciplined intuitions. However, a scientist would like explanations to be able to be proven wrong by means of empirical findings or by the conceptual rigour of a scientific theory. But these methods cannot directly reach the homology established by Jonas between the experience and teleology.

The resolution of this problem lies in the ongoing pragmatic circulation and mutual constraining between science and experience (Gallagher, 1997; van Gelder, 1999; Varela, 1996, 1999). Jonas has identified the initial step in suggesting that metabolism is at the root of intrinsic teleology and natural agency. He has also landmarked many other crucial phases in the development of value by following the thread of increased mediacy between organism and environment (essays 4, 6, and 7 in Jonas, 1966). He has described how perception/action and emotion co-arise in animal motility, and how sight and image-making bring on the possibility of objectivity and eidetic control of action as well as the capability of constructing a project of life when image-making is applied upon us. These are valuable signposts but they do not yet constitute explanations, least of all scientific ones. They are points of reference from which we can provisionally designate the phenomena to be explained but which will not be unaffected by further developments in theory-building, empirical findings, and experiential congruence. In such a circulation practically anything may be called into question, even the legitimacy of the phenomenological guidance used, if it leads to a greater, self-accrediting, coherence and pragmatic value. The proposal that autopoiesis provides a scientific account of the initial step of this circulation must therefore be taken very seriously.

### 3. Autopoiesis and teleology

It would be a major achievement if the theory of autopoiesis were capable of fulfilling the role proposed by Weber and Varela. It could directly engage Jonas's biophilosophy in a fruitful dialogue with research in biology, neuroscience, and cognitive science by providing a common locus of synthesis for problems that are normally treated separately. I will show that their argument points in the right direction, and puts some crucial elements on the table, but falls short of achieving its full potential because of their unspoken reliance on apparent implications of autopoiesis which are not elaborated in the theory and do not immediately follow from the original formulation. Spelling out precisely what these additional properties are and entail opens the door to answering other questions and formulating new distinctions not the least important of which is the relation between metabolism, agency and history. I will proceed here by examining the primary formulation of autopoietic theory and the distinct concepts of robustness and adaptivity and showing that autopoiesis implies only the former while sense-making needs the latter. Thus, autopoiesis *simpliciter* cannot provide the intended grounding for sense-making.

#### (a) *From autopoiesis to sense-making*

For a promising start, Weber and Varela suggest that Jonas's stipulation of pattern-forming tendencies in matter tallies well with theories of self-organization and non-linear phenomena of which he was probably unaware. It is encouraging that autopoiesis belongs to this broad set of self-organizing processes since some way of describing the mesoscopic behaviour of collective matter as active, as opposed to inert, was intuited as necessary by Jonas in order to account for the possibility of metabolism, (Jonas, 1979). An essential difference between autopoiesis and the rest of the wider class of self-organization is that what is by definition a process of material self-production must as a result generate a self-distinguishing concrete unity and not simply a physical pattern. The unity is self-distinguishing because it is constructed and sustained by its own activity in spite of the equalizing physical tendencies. This is now a step forward, for not only does autopoiesis describe much of what Jonas predicates of metabolism but it goes a bit further by making explicit the need for an actively constructed boundary that physically separates metabolism from the external medium.

As the next step, Weber and Varela propose that autonomous self-production leads to two complementary senses of teleology. One is the Kantian sense of mutual generative relations between organismic components and between them and the whole, making the living system a natural purpose. As we shall see, autopoiesis indeed provides a more sophisticated grounding for this intrinsic teleology. From here, they speedily describe self-production as a "concern to affirm life", (Weber & Varela, 2002, p. 116). This leads them to the second sense of teleology: "the instauration of a point of view" (ibid. p. 116) and an individuality that through self-production is "*ipso facto* a locus of sensation and agency", (ibid. p. 117). This projective teleology, or *sense-making* (see also Varela, 1991, 1997), would seem to follow naturally from the language of concerned self-affirmation, for an entity for whom its own continuation is an issue would immediately project this concern onto its surroundings.

If it could be shown that the original definition of autopoiesis directly supports this second sense of teleology, the work of Weber and Varela would be complete. If not, doubts would be cast about the relation between intrinsic teleology and sense-making for autopoiesis would be sufficient for one but not the other. The supporting logic of their proposal is similar to that of Jonas (see also Thompson, 2004): A. self-production is a process that defines a unity and a norm: to keep the unity going and distinct, B.

encounters with the external world are “evaluated” by the system (through the autopoietic machinery) as contributing or not to the maintenance of autopoiesis, consequently C. autopoiesis implies sense-making, an intrinsic perspective of value on the world. The crucial point here is B. and whether it can be derived from the original formulation of the theory.

(b) *Autopoiesis as conservation of autopoiesis*

The theory of autopoiesis is based on mechanistic assumptions and explicitly rejects teleology (Varela, 1979; Maturana & Varela, 1980). It proposes a distinction between two valid kinds of scientific discourse, the operational and the functional/symbolic, and bases itself on the first kind. Operational discourse belongs to the contemporaneous domain of physical processes operating in the living system (for instance, descriptions of physico-chemical or neural processes) and functional/symbolic statements are those formulated by an external observer given relational knowledge of the interaction and historical contexts (for instance, explanations of behaviour in terms of evolutionary advantages). This strict separation allows the theory to reach some of its boldest conclusions and justify them as implications of autopoiesis: cognizers cannot make use of internal representations; communication is not the exchange of information, etc. Much of the interest generated by autopoiesis is due to this strict systemic grounding of cognitive and biological terminology, so it should not be easily discarded. In itself, the rejection of teleological language is not a problem for the project of grounding sense-making. On the contrary, it is necessary for this grounding to work that the explanandum should not inadvertently make its way into the basic assumptions of the theory.

Central to the theory of autopoiesis is the axiom of structural determinism: changes of state in a system always operate in the present as a result of its current structure and are not determined by external agents or contextual conditions. With this axiom in mind, an autopoietic system is defined as

- a network of processes of production (synthesis and destruction) of components such that these components:
1. continuously regenerate and realize the network that produces them, and
  2. constitute the system as a distinguishable unity in the domain in which they exist, (Weber & Varela, 2002, p. 115)†.

Formally interpreted, this definition is insufficient to eliminate trivial exemplars and has often been supplemented with additional guidelines for its interpretation. For instance, autopoietic systems conserve their organization by “maintaining constant certain relations between components otherwise in continuous flow or change” (Maturana & Varela, 1980, p. 81), as opposed to systems that are static and simply endure. Autopoietic systems should be interpreted as being far from equilibrium and open to material and energetic exchange – though not implied by the definition, this interpretation is supported by the dynamic connotation of the word “process”. Another clarification is that by processes of production we should interpret *material* production in physical space (with the corollary that in order to make itself self-distinct the autopoietic system must build a physical semi-permeable membrane around itself)‡. Taking these interpretations on board, the self-bounded, self-regenerating material system already fulfils the Kantian

† Essentially equivalent definitions can be found in (Maturana & Varela, 1980, p. 79), (Varela, 1979, p. 13) and almost in every item of the primary literature. For a discussion on clarifying and extending this definition in ways connected with this paper, see (Bourgine & Stewart, 2004).

‡ The demand for physicality seems at first *ad hoc* and arbitrary in a definition that is proposed as strictly formal and systemic. This is an unresolved issue at the core of continuing debates about the possibility of computational, or other non-physical kinds of autopoietic processes (Boden, 1999; McMullin, 2004; Moreno & Ruiz-Mirazo, 1999). The existence of computer realizations of algorithmic

criterion for intrinsic teleology since the generative relations between processes define a closed network, and this closure means that single components can be interpreted naturally in both causal and finalistic terms in relation to other components and the whole.

It seems that some of the consequences of autopoiesis are not derived strictly speaking *ordine geometrico* but rely on appealing to intuitive notions to guide the interpretation of the terminology. This is generally not a major problem in so far as it shared by any scientific theory. In this context, how should we interpret the continuous realization and regeneration of processes in the autopoietic network? Two terms are used by Maturana and Varela: *conservation* and *homeostasis*. These terms have different connotations. Sometimes the emphasis is on homeostasis: an autopoietic system dynamically maintains certain relations invariant through compensatory structural changes and the homeostatic variable is the system's own organization (Varela, 1979, p. 13), (Maturana & Varela, 1980, p. 79). Other times the emphasis is on conservation as opposed to change (Maturana, 1975). This is a more abstract and neutral notion.

Whereas homeostasis connotes the existence of active mechanisms capable of managing and controlling the network of processes that construct the organism, conservation is a set-theoretic condition that may or may not be realized in an active manner. It merely distinguishes between changes of state without loss of organization and disintegrative changes. In typical autopoietic parlance, the organization of a system is conserved as long as it enters in relations with its environment that do not affect its autopoiesis. "Otherwise it disintegrates" is the phrase that often follows in the primary literature, (e.g., Maturana & Varela, 1980, p. 98). Frequent moves like this do a clean sweep with the possibility of any teleological interpretation that the idea of homeostasis may allow. It is not because the organism has struggled to conserve itself that we can observe it now, it is because we see it now that it must have remained viable up to this point. Accordingly, there need not be any tendency originating in the organism to counteract *potential* loss of viability (in this view, that would be like a structure-determined system acting in response to a future event and thus against structural determinism).

The relation between autopoiesis and sense-making hinges on where the emphasis is put: abstract conservation or active homeostasis. It soon becomes clear that autopoiesis, a concept allowing no degrees, can only entail the more general case of conservation (thus admitting the possibility of fortuitous self-maintenance) and not the special case of active homeostasis. Recently, Maturana has described conservation of organization and adaptation as the constitutive conditions for the realization of a living system (Maturana, 2002, p. 11). This consequence of the original definition is in tension with Varela's later views of autopoiesis as a self-affirming process of identity because conservation, in as much as it is able to suggest a norm for sense-making, can suggest only an all-or-nothing norm: organisms live as long as they don't die. Tautological though it sounds, thanks to the definition of autopoiesis, the content of this statement is not void since it allows the

processes showing sometimes surprisingly lifelike properties such as growth, membrane construction, and self-repair (e.g., Ono & Ikegami, 2000) belies the fact that removing physicality from autopoiesis elevates trivial examples such as attractor states in cellular automata to the category of the living. Varela (1979) attempts to resolve the issue by defining *autonomy* or generalized operational closure as the class of self-producing processes occurring in arbitrary domains. Autopoiesis is the restriction of this class to the physical, molecular domain. But this merely defers the problem which is now that of trivial cases of autonomy. It is very likely that physicality brings more to the interpretation of autopoiesis than initially anticipated and that these elements also need to be spelled out for an improved definition of life. This worry had led some researchers like Alvaro Moreno and colleagues to explore the implications of thermodynamic constraints on autopoiesis, (see Ruiz-Mirazo & Moreno, 2004). The important result that this line of research is achieving is precisely determining what it is about physical constraints that make real life non-trivial. Interestingly, anyone who believes that life is a formal, substrate-independent, property, as opposed to formal *and* material, should follow these developments closely because any chance of substantiating their beliefs lies in being able to formalize the restrictions brought by physical constraints that make life a non-trivial affair.

observer to distinguish between lethal and non-lethal encounters based on the particular mode of achieving self-production. It is sufficient for point A above, but not for point B which would require the autopoietic system itself to be able to make the same distinction†.

According to the conservation perspective, balancing at the edge of a cliff is a perfectly viable behaviour, so is falling over, both are interactions that conserve autopoiesis, it is only crashing against the ground that is bad for the organism. Such all-or-nothing normativity is clearly stated in the definition and use of the concept of *structural coupling*, i.e., encounters with the environment resulting in perturbations to the autopoietic dynamics without loss of organization (used throughout the literature, but first defined in Maturana, 1975). This important concept is the basis of all that autopoietic theory has to say about cognition and a key element in the enactive approach (Varela, Thompson, & Rosch, 1991). And yet, if any value judgment could be formally derived from structural coupling it would be the rather useless *a posteriori* realization by the external observer that the organism should have avoided that very last encounter that killed it – it would be, unsurprisingly, no sense-making at all.

Something other than the set-theoretic notion of conservation must be what Varela has in mind when he speaks of a perspective on the world that originates in the very same process of organismic identity and which is haunted by a *surplus of signification*, (Varela, 1997). Or when he emphasizes that signification is always the result of a *lack*, something that is missing from the recurrent process of self-production, and therefore meaning must originate in *breakdowns* of autopoiesis, be they minor or major, (Varela, 1991, pp. 86–87).

The conservation view should lead to a stern refusal of such language. There is no room for concepts such as lacks, minor or major breakdowns in autopoiesis: either organization is conserved or it isn't – being partially autopoietic is senseless and any notion of the system being *at risk* of disintegrating would be a remark made by the external observer and plays no operational role. Similarly, the original formulation implies that the system does not adapt to changes in its environment, it is already adapted to it and either survives the perturbations or it doesn't, thus rendering the notion of adaptation useless (incidentally leaving no room for important biological phenomena such as stress, illness, or fatigue). For Varela's argument to work while remaining compatible with the theory of autopoiesis there must be some additional element, some particular way of realizing autopoiesis that admits of graded notions such as lacks and breakdowns and articulates in detail how signification is generated.

Consider the example used by Varela on several occasions: bacteria swimming up a sucrose gradient, (Varela, 1991, 1997). Accordingly, the autopoietic network selects some particular physical correlations (in this case the presence of sugar) as relevant. Undoubtedly, it is the self-producing organization that sets sugar apart from other chemical compounds. This is in the first instance a case of the environment being relative to the

† The situation is remarkably similar to the problems surrounding Spinoza's doctrine of self-preservation. Although Spinoza makes use of his argument that a thing's essence is its endeavour (*conatus*) to persist in its being (*Ethics*, IIP4-7) to generate a system of values, the argument is fraught with problems originating in Spinoza's avoidance of teleology. For him, *conatus* belongs to the general concept of appetite which he analyses non-teleologically in terms of tendencies to bring about certain events, such as the hungry organism having a tendency to get closer to food. But this interpretation suggests a weak version of self-preservation: simply that a thing's nature will not make it act so as to obtain its self-destruction (IIP4). This statement which may be false for things in general (Matson, 1977) is indeed true of autopoietic systems because their essence is self-production. Nevertheless, even if no autopoietic system can initiate its own destruction, we cannot deduce therefrom that it seeks self-preservation in the active, finalistic sense. Autopoiesis may produce an effect that will help autopoiesis, but it will not necessarily produce an effect *because* it will help autopoiesis, see (Bennett, 1984, 1990; Curley, 1990). More on Spinoza's potential for a theory of the organism can be found in (Jonas, 1979) and (Duchesneau, 1974).

organism<sup>†</sup>. Moreover, for bacteria, sugar is an essential *nutrient* that keeps the bounded network of processes of production going. Thus bacterial autopoiesis is like an arrow that points *us* to sugar as a particularly relevant physical factor out of many.

But what makes bacteria *swim up* the gradient? What makes *them* distinguish and prefer higher sugar concentrations? As defined, structural coupling is a conservative, not an improving process; it admits no possible gradation. If the concentration is enough to keep bacteria viable the latter should be equally – not more – viable in a range of higher concentrations. Even if the current rate of nutrient intake is lower than the rate of consumption (leading to certain loss of autopoiesis in the near future), bacteria will not seek higher concentrations *just because* they are autopoietic since improving the conditions of self-production is not part of the definition of autopoiesis. Only if they are able to monitor and regulate their internal processes so that they can generate the necessary responses anticipating internal tendencies will they also be able to appreciate graded differences between otherwise equally viable states. Bacteria possessing this capability will be able to generate a normativity *within* their current set of viability conditions and *for themselves*. They will be capable of appreciating not just sugar as nutritive, but the direction where the concentration grows as useful, and swimming in that direction as the right thing to do in some circumstances.

### (c) *Adaptivity*

What we need to reconstitute Varela's remarks about sense-making is to articulate the above capability so as to ground breakdowns in operational terms. To do so it is convenient to put in different terms the contrast between what autopoiesis implies and what sense-making requires. Autopoietic systems exist far from equilibrium and must tolerate the natural entropic trends by remaining energetically and materially open. Being tolerant is another way of saying that they are *robust* i.e., that they can sustain a certain range of perturbations as well as a certain range of internal structural changes before they lose their autopoiesis. These ranges are defined by the organization and current state of the system and are here referred to as its viability set which is assumed to be of finite measure, bounded, and possibly time-varying. Robustness implies endurance but not necessarily *adaptivity* which is a special manner of being tolerant to challenges by actively monitoring perturbations and compensating for their tendencies.

Adaptivity is defined as:

a system's capacity, in some circumstances, to regulate its states and its relation to the environment with the result that, if the states are sufficiently close to the boundary of viability,

1. tendencies are distinguished and acted upon depending on whether the states will approach or recede from the boundary and, as a consequence,
2. tendencies of the first kind are moved closer to or transformed into tendencies of the second and so future states are prevented from reaching the boundary with an outward velocity.

This capacity may result from the action of dedicated mechanisms or it may be an emergent aspect of specific ways of realizing autopoiesis. But importantly, it is not a direct consequence of the definition of autopoiesis but an elaboration of it from which it

<sup>†</sup> Such is the case for *any* system, autopoietic or not, if the environment is defined as the set of external variables that can affect the system, and the set of parameters for external processes that are affected by it, (Ashby, 1960, p. 86). The organization of the thermostat points to changes in local temperature as the relevant environmental variable as opposed to, say, the level of sound. What is special about autopoietic systems is that some of the environmental encounters are essential for their continued construction. Bacteria are partly made of sugar; thermostats are not made of temperature changes.



is possible to recover the homeostatic interpretation. A breakdown is simply the severity of a negative tendency distinguished and measured by the amount of regulative resources that it demands to compensate for it with or without plastic re-structuring of the system. A breakdown will typically, but not exclusively, be the result of external perturbations, and in addition to responding to them, adaptivity allows organisms to avoid some risky situations and seek preferable ones.

Only of the subclass of autopoietic systems that are not just robust but also adaptive can we say that they possess enough operational mechanisms to distinguish the different implications of equally viable paths of encounters with the environment<sup>†</sup>. If sense-making requires the acquisition of “a valence which is dual at its basis: attraction or rejection, approach or escape” (Weber & Varela, 2002, p. 117), a sense-making system requires, apart from the norm given by self-construction, access to how it currently stands against the all-or-nothing barrier given by that norm. In order to have such access while operating as a consequence of contemporaneous states, an autopoietic system must be able to recognize in those states, and only in them, the tendencies that relate it as a whole to the potential loss of its own viability *and*, in addition, it must be able to act appropriately on those tendencies.

Both elements, self-monitoring and appropriate regulation, are necessary to be able to speak of meaning from the perspective of the organism. Self-monitoring without the right response is (apart from useless) meaningless, since significance must relate to a referential totality (in this case a totality of operations internal to the system). This *paradigmatic* aspect of meaning is provided by the actions of the counteracting mechanisms which differ in degree or in kind for different encounters. Events that provoke the same regulative response are not meaningfully distinguishable. Similarly, regulation without the guidance of self-monitoring is (apart from not possible) disconnected from the source of *syntagmatic* meaning that links the right concatenation of responses to the right situation, that is, its neutralizing or ameliorating effect and subsequent evaluation as such.

We find that sense-making in organisms needs both autopoiesis and adaptivity. Autopoiesis provides a self-distinct physical system that can be the centre of a perspective

<sup>†</sup> This subclass would seem to cover all known organisms. Indeed, adaptive autopoiesis is closely related to what Margaret Boden (1999) describes as the strongest and most widespread sense of metabolism which involves energy budgeting and mechanisms capable of selecting the most appropriate bioenergetic path according to the context. Nevertheless, robustness without adaptivity is far from being an exotic and empirically empty possibility. It is clearly the case in practically all models of minimal autopoiesis (Varela, Maturana, & Uribe, 1974; McMullin, 2004; Bourguin & Stewart, 2004) as well as in the current examples of synthetic chemical autopoiesis (Bachmann, Luisi, & Lang, 1992; Szostak, Bartel, & Luisi, 2001). Typically these systems consist of a metabolic reaction transforming external nutrients into both the elements necessary for the continuation of the reaction and the construction of the membrane. Through self-organizing mechanisms (e.g., vesicle formation), the membrane encloses enough concentration of reactants so that the system produces itself as long as the external conditions are within viable ranges. There is no explicit or emergent monitoring function in such models. Whenever a link in the network of processes is broken, if the remaining processes are fast enough to repair it, they do so by having the same function they had before the breakdown. The system tolerates such perturbations, it does not adapt to them. Adaptivity can be added to such minimal systems in many ways and to different degrees. For instance, Tibor Gánti (2003) in the 1970s independently proposed a theory of minimal living systems which has many parallels with autopoiesis. The Chemoton, the accompanying model to this theory, includes all the mechanisms of autopoiesis. It also includes an explicit cycle of template-based polymer replication which becomes active only above a certain threshold of concentration of internal products. The cycle uses these products to build some of the components that later will form part of the membrane. The overall result is that the system is able to respond to some increments in internal pressure (a global variable that must be regulated if the cell is to avoid exploding) by contingently constructing more membrane elements (and so growing in size) until the cell has grown so large that it is energetically more convenient to divide into two cells. This example shows that adaptivity is a perfectly operational property, and that measurement and regulation do not require any homuncular agency at play.

on the world, and a self-maintained, precarious network of processes that generates an either-or normative condition. Adaptivity allows the system to appreciate its encounters with respect to this condition, its own death, in a graded and relational manner while it is still alive. I suspect that some version of adaptivity was probably implicitly granted by Weber and Varela to autopoietic systems, particularly if emphasis is put on the homeostatic interpretation. But this emphasis is in need of articulation and the concept of adaptivity performs that job. What remains to be seen is whether there is more to adaptivity than a mere clarification.

#### 4. Adaptivity, agency, and time

Adaptivity allows an organism access to the implications of the mutually causal links between the processes that achieve self-production; this teleology can otherwise be appreciated only by an external observer and cannot yet be sense-making. The dialectics of monitoring and regulation, poles which may be indistinguishable in terms of mechanism because they are in fact co-defined, allow meaning to be generated by the system and for the system. It is important to notice that adaptivity cannot do any of this in the absence of self-production. Here, Jonas's criticisms of cybernetics still apply, (Jonas, 1966). A feedback mechanism, for instance a closed sensorimotor loop, may instantiate a purpose, but this purpose is external to it, the mechanism by itself does not generate it. In contrast, a norm is generated by autopoiesis, the natural distinction between self-maintenance and disintegration, but this distinction is not yet accessible to the autopoietic system unless it is also able to regulate itself with respect to this norm<sup>†</sup>. How and when this capability appears in the history of life is beyond the scope of this article, but it is not difficult to conceive a set of possible explanations ranging from the evolutionary to the dynamical (e.g., Ashby, 1960).

We may briefly enumerate a few interesting aspects of adaptivity as defined in the previous section. Firstly, it does not imply any kind of optimal or ideal state for the organism. Regulation may be active only in certain regions of the viability set leaving the possibility of neutral spaces where the system is, in the absence of some other norm-generating process, equally adapted. Nevertheless, adaptivity establishes an inner norm, a distinction between good, neutral, and bad ways of realizing autopoiesis. This is not disconnected from a second aspect: adaptivity is a graded property. Regulation may function over a large range of conditions or only in selected extreme situations.

Adaptivity may also be recursive in two connected senses. Firstly, adaptive processes may regulate states not in terms of their proximity to the viability limits, but depending on how close they are to activating more basic adaptive responses, thus permitting the meaningful distinction of events that do not put the system directly in any danger. For instance, the preference for more nutritive food may well be regulated by mechanisms that adaptively avoid the activation of more serious bioenergetic regulation (such as the consumption of reserves) which itself inherits a negative valence from its proximity to the boundary of viability. In a second sense of recursiveness, regulative processes may themselves be plastically adjusted or transformed depending on their effectiveness. Here again, the norm that guides such higher levels of regulation is internally generated by the underlying adaptive processes.

One further important aspect of this emerging picture is that only thanks to adaptivity can we speak of organismic dysfunction, stress, fatigue, mal-adaptation, and pathology. Autopoiesis in the conservation view is blind to such phenomena since they all occur

<sup>†</sup> If intrinsic teleology does not yet guarantee sense-making, how does this affect Jonas's existential argument about the ontology of metabolism? It leaves it intact, for this argument moves in the opposite direction, from the embodied experience of sense-making to the intrinsic teleology of metabolism and this entailment is maintained in the present analysis.

while the system is still autopoietic, but adaptivity provides a measure for them. Indeed, it is possible to define these phenomena in terms of failures of adaptivity such as the exhaustion of adaptive resources, malfunction of regulation, loss of adaptive buffering provoking the activation of extreme regulation, disharmonious activation of conflicting adaptive mechanisms, and so on. Thus, by re-establishing an adapted state, possibly through the simultaneous repair of adaptive processes and change in the range and kind of acceptable relations with the environment, a successful cure may well *re-define* rather than simply *restore* the organism's own normativity; health, from this perspective, is very different from a statistical species-specific correlation of normality and there are consequently many ways of being healthy, (Canguilhem, 1966; Goldstein, 1934).

All these aspects deserve more extensive development. Here, however, I will focus only on three important points that complement the preceding discussion. The first is a distinction within what has, so far quite generally, been called sense-making that is better articulated once the concept of adaptivity is firmly established. The second is a definition of agency and behaviour different from those given in autopoietic theory. The third concerns the temporal structure of adaptivity and sense-making.

(a) *The construction of sense*

The term *sense-making*, used by Weber and Varela to mean the enjoyment of a concerned perspective on the world, has two connotations. The first is emphasized by the everyday phrase "to make sense of something", i.e., to reveal or uncover its established meaning. The second is highlighted by the active use of the verb "to make", i.e., to generate or construct new meaning (for instance, by learning new correlations). These correspond rather well to different aspects of adaptivity: the regulation provided by established adaptive processes and the plastic modification, attunement or generation of novel adaptive processes respectively. Encounters between the organism and its environment derive any meaning they already have from the corresponding activation of regulatory mechanisms that assimilate them into an existing *form* within the autopoietic organization.

This form emerges from the dynamical preferences that the whole internal adaptive landscape defines within the confines of viability: the organism's own particular way of realizing and regulating autopoiesis. But, as already suggested, the internal landscape need not be a static one. Its form is constrained but underdetermined by autopoiesis. The possibility is therefore open for plastic mechanisms that regulate adaptivity to adjust the organism to novel regularities in the set of those initially neutral encounters and to accommodate the current structure of meanings by replacing previous correlations with new ones.

We find that the distinction between *uncovering* and *constructing* meaning is not possible if autopoiesis is taken to be all that is necessary for sense-making. The autopoietic organization may well select an external encounter, e.g., sugar, as relevant for the ongoing renewal of its network of processes, but that alone cannot explain why an organism would ever give meaning to any encounter that is not itself *physically* involved in autopoiesis and yet can act as a proxy for something else that is, e.g., the smell of food. For this, the capability of meaning-*construction* must be in place.

Even if the arguments in the preceding section were wrong and autopoiesis were enough to ground sense-making, it would still be confronted with the problem that the only sense-making that could be so grounded would be the *ipso facto* uncovering of *direct* relevance of an external encounter for continued self-production. The problem of how meaning can arise *ex post facto*, through durable processes of assimilation and accommodation of novel, initially neutral but retroactively significant encounters would remain unsolved, and so would the problem of how encounters with no immediate physical

implications for autopoiesis derive their meaning as being the reliable bearers of future consequences for viability.

Adaptivity solves these problems. Meaning-construction is enabled by the plastic attunement and re-organization of the web of adaptive processes in an experience-dependent manner thus producing through the organism's ontogeny new parameters to observe and new responses to their changes. And because adaptivity works mainly on tendencies of states, it needs not wait until there is a direct contact between the organism and the autopoietically relevant encounter to start responding to its potential presence in an anticipatory manner or recursively to other bearers of meaning. Snowprints left by a prey are meaningful even though metabolism is not feeding on the prints themselves.

(b) *Natural agency*

Let us now turn to the problem of agency. Adaptivity, so far described as a set of internal regulatory mechanisms, can indeed operate at the boundary of the organism and beyond. Parametrical regulation such as active transport through the selective opening and closing of ion channels is widespread in uni-cellular organisms and is one of the most common examples of control of the conditions of physical exchange between organism and environment. More sophisticated control involves the whole cell, as in the displacement towards nutrient-rich regions of the medium, or the construction of protective biofilms, or the projection of pseudopodia to engulf another cell in phagocytosis. Notice that the nature of the physical coupling is always fully dictated by the laws of physics, a cell cannot simply change the laws of reaction and diffusion. But what is given to the organism is the parametrical control of those laws by its influence on the constraints of the coupling dynamics.

This observation allows us to draw the important distinction between structural coupling and the *regulation* of structural coupling. The former is an ongoing happening, the necessary outcome of non-lethal physical encounters between organism and medium. Only the latter, the parametrical action that regulates coupling, fully deserves the name of *behaviour* because such regulation is *done* by the organism – even if it does not always lead to an improved condition – as opposed to simply being *undergone* by it. Unregulated coupling is better described as suffering an exchange while behaviour is the control and selection of what exchanges to suffer.

Autopoietic theory makes no such distinction and simply defines behaviour as the structural changes an organism undergoes while maintaining its autopoiesis (Varela, 1979, p. 240), (Maturana & Varela, 1980, p. 124) and that are describable as movements or changes of attitude in relation to the environment. Receiving a blow or being stung by a wasp are accordingly as good examples of behaviours as punching or stinging. The problem is not that this definition of behaviour goes against the use of the term in every day life and scientific discourse. The problem is that a distinction which is crucial for cognition is thus lost, a distinction which paradoxically is what Weber and Varela seem to be trying to recover in the project of naturalizing sense-making. Cognition requires a natural centre of *activity* on the world as well a natural *perspective* on it. While Weber and Varela mainly speak of the second, it is clear from the literature on the enactive approach (Varela et al., 1991) that the first is as important to them. However, structural coupling alone cannot ground the concept of activity any more than autopoiesis alone was able to ground the concept of perspective, and for the same reasons.

Activity, like perspective, is an asymmetrical concept. There is the actor and that which is acted upon. But just because autopoiesis establishes a self-distinct physical unity doesn't mean that the necessary asymmetry has been achieved in the domains of exchanges between the unity and its medium. Structural coupling refers to the mutual perturbation between organism and environment and this exchange may or may not

subserve a tendency towards mediate conservation of autopoiesis. Only when a process is established that is able to regulate this exchange so that in general the result is an improved condition of viability, only then it is possible to speak of a true asymmetry. Regulation is done by the organism and for itself; there is no analogous process in the general case originating in the environment. Behaviour defined not as physical coupling, but as its regulation, is always asymmetrical, has an intentional structure, and can be said to either succeed or fail. It is only at this stage, when the organism behaves, that we may speak of an *agent* (following a similar use advocated by Ruiz-Mirazo & Moreno, 2000), i.e., a self-constructed unity that engages the world by actively regulating its exchanges with it for adaptive purposes that are meant to serve its continued viability<sup>†</sup>.

A consequence of this definition of agency which is relevant for the project of artificial intelligence is that the intentional structure of behaviour is inevitably related to both what the organism is and what it is likely to become, even when it may also relate to other things (such as its own form, and its connection to other acts). The movement of meaningful action can be convincingly emulated in an artificial system but this is not the same as the system acting meaningfully. The robot may look scared and retreat when yelled at but this may be only a sophisticated illusion; there is no way to tell just from observing its responses, Turing-test style. Being functionally scared is not the same as being scared. To check this operationally is relatively simple (at least to rule out negative cases). We only need to look into the organization of the system and reveal its self-generated structure of sense-making (if any) and its relation to the system's viability. The vast majority of current work on robot emotion will fail this test dismally due to its almost exclusive reliance on pure functionality. Here I shall leave the question open as to whether such artificial intentionality is possible short of reproducing a fully metabolizing robot that is an agent for itself and not by fiat. The conceptual landscape changes once non-metabolic values enter the picture.

### (c) *Temporality*

A final point regarding adaptivity is that it brings to the organism its own multi-scaled and directed temporality and the enabling of a historical dimension. Adaptive regulation can properly be called an *act* in the general sense given by Langer (1967). It is a structured event, with clearly defined phases of onset (the sensing of a negative tendency), acceleration (the activation of the adaptive mechanism), consummation (the overturning of the negative tendency) and cadence (the de-activation of the adaptive response). These phases are reliably, if not invariably, present and distinct. The form of adaptivity is such that a given stage always pre-shapes the next one without fully specifying it. Adaptive events thus have a *temporal direction* that autopoiesis (surprisingly) lacks. Being a conserved quantity, autopoiesis is also conserved if we invert the flow of time (the network of inverted reactions still maintains itself, waste products and nutrients exchange roles), but adaptivity (unlike robustness) becomes dysfunctional by converting safe conditions into dangerous ones which are, in this thought-experiment, fortuitously nullified by environmental encounters.

The adaptive event (or act) may be formed by the concatenation and parallel co-ordination of many other regulatory events, but there will be a point below which no further de-composition will be possible without losing the time-structure of the act. At

<sup>†</sup> As adaptivity is generally not implied by autopoiesis – there may be autopoietic systems that survive without it – so agency is not implied by autopoiesis and adaptivity combined. There can be adaptive autopoietic systems where regulation is circumscribed to internal responses to external encounters without any active regulation of the conditions that affect these encounters (the Chemoton would model such a system). There can also be different degrees of agency measured by the organism's capability to control and alter its body and environment. The transition to animality discussed by Jonas is in this view a jump in the degree of organismic agency.

that point what remains are raw processes<sup>†</sup>. There is consequently a *minimum temporal granularity* in adaptivity. By way of example, it is not possible to judge if the sudden overproduction of a metabolite is part of an adaptive response unless the analysis is extended to a minimal period of observation spanning the immediate past and future.

History follows from the granularity and the time-asymmetry introduced by adaptivity: the possibility of neutral valence means that certain regulative responses may compensate negative tendencies and leave the organism viable but changed, i.e., marked by the encounter in ways that constrain further dynamics. Those changes may then be reflected in the working of further regulation. The effects are propagated, eliminated, transformed or generally time-managed by sequential or hierarchical regulative events. A *historical dimension* is thus inaugurated which is not merely the contingent reflection of environmental encounters nor is it governed entirely by an internally generated plan.

The directed, historical, and granular temporality of adaptivity is reflected in the time-structure of sense-making in a way that squares well with embodied experience which also shows the same aspects of pointedness, history, and minimal time-span. The operation of single adaptive mechanisms is in normal circumstances self-extinguishing but their interaction, the ongoing coupling with the environment, and the precariousness of metabolism, make their collective action also self-renewing, thus naturally resulting in *valenced rhythms of tension and satisfaction*. This is yet another temporal aspect of our own experience of sense-making (in our case realized through adaptive processes involving the whole body and the nervous system) that would remain mysterious were it not for adaptivity.

## 5. The road ahead: non-metabolic values and dynamic forms

Weber and Varela close their article with a forward-looking and, in their words, immodest conclusion: a science of biological individuality and teleology is possible in the shape outlined by the theory of autopoiesis. This paper can only agree with this vision and should be seen as an effort to unpick the threads of this proposal and solve the problems that become apparent in the process. If autopoiesis in the present analysis suffices for generating a natural purpose, adaptivity reflects the organism's capability – necessary for sense-making – of evaluating the needs and expanding the means towards that purpose. I will conclude my critical remarks with even more immodest speculations that highlight how much there is still to be done even after disclosing a good fraction of the mileage provided by the elastic concept of adaptivity.

We have now a clearer understanding of the minimal processes involved in intrinsic teleology, sense-making, and agency. This formulation prompts the questioning of the nature of values that *do not* originate in the norm established by metabolism but elsewhere: in self-sustaining, operationally closed processes that work within the boundaries of metabolic viability but are underdetermined by it and can consequently introduce their own normativity. Examples include motor intentions within organized patterns of sensorimotor coordination, socially constructed values, and human projects.

Jonas explores the development of the inner world of organisms by following the unfolding of the dimension of mediacy and freedom. A new order of values is found in animality with the arrival of motility and the co-emergence of perception, action and emotion. A transition towards objectivity is encountered in some uses of sight and

<sup>†</sup> Curiously, Langer doesn't seem to share this conclusion even if her theory already establishes many of the points made in this paper. In talking about the distinction between (behavioural) acts and movements, she calls this distinction artificial (Langer, 1967, p. 266). But, if an infinite regression of time-structured acts were allowed to exist in the composition of any act, the whole concept would be indistinguishable from that of a mere physical event. Here I am not suggesting that such events are some impossible kind of formless hyle, but simply that their form is no longer or not yet that of an act. Otherwise all events are acts and the concept loses its usefulness.

later in the capability of image-making, eidetic control and eidetic projection, so that behaviour can be guided by an as yet inexistent state of affairs and not only in response to the current situation. These transitions, Jonas suggests, are irrevocable because they introduce a new set of values, or more properly, a novel way of generating values. Why exactly this is so, Jonas doesn't expand on. But there would be a tension if he were to tie each of these new ways of generating values always back to the logic of metabolism.

Jonas recognizes that there is no metabolic gain in more freedom at the cost of more precariousness, such as in the transition to animality. If motility results in the balancing of the same equation with larger positive and negative terms, there is no reason why it should not be able to evolve back and remove its supposed gains in exchange for increased safety. Jonas intuits that this cannot happen (except for the extinction of animals). But if this is true, then it must be because a *new form of life* is born in animality (and in image-making and in other transitions). This new form of life is not contrary or indifferent to metabolism. Made possible by the latter, it will mostly be at its service, but it may also be independent of it to the extent that the adaptivity of metabolism does not dictate a unique way of doing what is necessary for its continuation. Within this independence, the new form of life will be able to generate, via a process of adaptive closure analogous to metabolism, its own set of values, thus making the process irrevocable and resulting in the *coexistence of different identities in a same organism*. This is why animal action has its own organization, a specific preference out of many metabolically compatible options; it is also the reason why the processes that maintain such organization (occurring in patterns of nervous activity which are already one level removed from direct metabolism) themselves can introduce a norm in potentially constant development according to which the action acquires intrinsic value on top of its functionality: the dexterous movement, the awkward posture, the confident gesture, the elegant walk.

In order to explain these preferred modes of achieving a metabolically prescribed end (e.g., self-protection, foraging), it is sufficient to posit a similar kind of self-sustaining, self-generating dynamic form in animal behaviour and in neural and bodily activity which is reflected in postural habits, perceptual invariants and organized action. The idea of *form* plays a fundamental role in this same context in the work of Goldstein (1934), Langer (1967), and Merleau-Ponty (1962, 1963). The latter's concept of motor intentionality provides the clearest parallel to the preceding discussion on the structure of adaptivity. If through a history of coupling between sensorimotor coordination and plastic regulation affecting the neural, bodily, and environmental factors that enable action, a regular form of behaviour is reached, this will tend to be conserved in so far as it involves adaptive processes that regulate its maintenance. The consequence of this is that the elements of a behavioural act, apart from their causal links, acquire an intentional structure that bi-directionally connects one naturally to the next and all to the act as a whole. Repair can then follow a breakdown in organized action even if the breakdown does not directly affect the self-affirmation of metabolism.

Thus it is possible to understand, for instance, how human vision can adapt to wearing bi-coloured or prismatic goggles (Kohler, 1964), a kind of sensorimotor disruption that cannot conceivably be thought of as placing the organism in any *direct* metabolic risk. It also becomes clear why the process of adaptation in such cases occurs only after the repeated activation of the behavioural patterns that generate the necessary value to distinguish between success and failure of an act. Metabolic adaptivity should be able to achieve the same result even if the subject were sitting still and waiting. But this is not what happens; re-adaptation only occurs if the self-affirming patterns of perception and action are given a chance to discover that something has gone wrong. Such patterns can also explain the self-generated value of compulsive behavioural cycles, such as substance addiction, that may even work *against* metabolic viability. Without this possibility, that is, if all organismic values were the result of a unique value-generating process, i.e., a

unique organismic identity, inner conflict would never arise and every disharmony of values experienced by an individual would originate only in the clash with the values of others. The clearest proof that such non-metabolic values are real and sometimes in conflict with the logic of metabolism is given by the experience of undergoing a prolonged and often difficult process of re-habilitation in embodied patterns of habitual behaviour. Formulating a theory of the organization of complex activity and its value-generating properties is now a major challenge for cognitive science and bio-inspired AI.

Experience also points to the existence of values and purposes not directly linked to the norms obtainable either from metabolism or the logic of self-sustaining complex behavioural forms. Those are values such as care of the offspring or sexual valence which, even though they make perfect evolutionary sense, do not seem to be associated with the self-maintenance of any process in the here and now. Jonas's most radical idea is his suggestion that all value originates in precarious and contemporaneous processes of self-constructed identity. They are in a very real sense something that organisms *do*, and not an abstract explanatory principle. If this is true, then the implication of evolutionary values not linked to the logic of metabolism is that somehow value-generating processes can be *shaped* to orient or extend valence into evolutionarily functional purposes.

As much as birds need real wings for flying, values that serve a purpose linked to the reproductive advantage of organisms, their kin or their group, must result from contemporaneous processes. Little is explained by terms such as *instincts* or *drives* if we cannot also show what are the underlying value-generating processes involved. It would be equivalent to positing the concept of abstract flight-enabling structures and not understanding what it is about wings that allows birds to fly. This is the radical shift that comes from taking Jonas seriously, for evolutionary values cease to be part of a valence-agnostic norm that emerges from a population-wide cost/benefit analysis, but are in fact embodied in the organism. Consequently, they must be the result of value-generating processes such as metabolism or closed, self-sustaining dynamic forms, whose valence is affected by selective processes to reflect the evolutionarily relevant norm (for example, care of the offspring or niche defense could be extensions to care of the self).

But there may be a deeper answer to this problem stemming from a critical analysis of biological reproduction as a temporally extended process. Autopoiesis rightly proposes self-production as ontologically prior to reproduction. However, once established, it would be mistaken to think of reproduction as a further trick added on top of metabolism. The organism's historical dimension serves as a nexus between the indefatigable continuity of metabolism and the punctuation of reproductive events so that they need not be seen as disjoint processes. The situation is made only more ambiguous in animals where there is a shared metabolic relation between mother and offspring during gestation that is itself an extended and formative process. Even if reproduction can transcend biological individuals, it still partakes of metabolism and there is no principled reason why it cannot generate values of its own for those individuals.

We can only be speculative at this stage about the processes underlying non-metabolic teleology and the general malleability of value-generation that takes place during evolution or indeed in the "engineering of the self" that is manifested in socially mediated human projects. The important point is that the potentially workable form of these proposals follows directly from the shape of Jonas's biophilosophy proving once again its far-reaching implications.

### Acknowledgements

Thanks for their sensible suggestions to Xabier Barandiaran, Alvaro Moreno, Marieke Rohde, John Stewart, and Steve Torrance. Special thanks to Andreas Weber for his extensive comments, Arantza Etxeberria for pointing out some issues about history and reproduction, Inman Harvey for suggesting an improved definition of adaptivity, and Maggie Boden for her probing points about cats, penguins, oak trees in distress, and angels.



## References

- Ashby, W. R. (1960). *Design for a brain: The origin of adaptive behaviour* (Second edition). London: Chapman and Hall.
- Bachmann, P. A., Luisi, P. L., & Lang, J. (1992). Autocatalytic self-replicating micelles as models for prebiotic structures. *Nature*, **357**, 57 – 59.
- Bennett, J. (1984). *A study of Spinoza's Ethics*. Cambridge University Press.
- Bennett, J. (1990). Spinoza and teleology: A reply to Curley. In Curley, E., & Moreau, P.-F. (Eds.), *Spinoza, issues and directions*, pp. 53 – 57. Leiden: E. J. Brill.
- Boden, M. A. (1999). Is metabolism necessary?. *Brit. J. Phil. Sci.*, **50**, 231 – 248.
- Bourgine, P., & Stewart, J. (2004). Autopoiesis and cognition. *Artificial Life*, **10**, 327 – 346.
- Canguilhem, G. (1991/1966). *The normal and the pathological*. New York: Zone Books.
- Curley, E. (1990). On Bennett's Spinoza: The issue of teleology. In Curley, E., & Moreau, P.-F. (Eds.), *Spinoza, issues and directions*, pp. 39 – 52. Leiden: E. J. Brill.
- Duchesneau, F. (1974). Du modèle cartésien au modèle spinoziste de l'être vivant. *Canadian Journal of Philosophy*, **3**, 539 – 562.
- Gallagher, S. (1997). Mutual enlightenment: Recent phenomenology in cognitive science. *J. Consc. Studies*, **4**, 195 – 215.
- Gánti, T. (2003). *The principles of life*. Oxford: Oxford University Press.
- Goldstein, K. (1995/1934). *The organism*. New York: Zone Books.
- Jonas, H. (1966). *The phenomenon of life: Towards a philosophical biology*. Evanston, IL: Northwestern University Press.
- Jonas, H. (1968). Biological foundations of individuality. *International Philosophical Quarterly*, **8**, 231 – 251.
- Jonas, H. (1979). Spinoza and the theory of organism. In Grene, M. (Ed.), *Spinoza: A collection of critical essays*, pp. 259 – 278. Notre Dame, IN: The University of Notre Dame Press.
- Kohler, I. (1964). The formation and transformation of the perceptual world. *Psychological Issues*, **3**, 1 – 173.
- Langer, S. K. (1967). *Mind: An essay on human feeling. Volume 1*. Baltimore, USA: The Johns Hopkins Press.
- Matson, W. (1977). Death and destruction in Spinoza's Ethics. *Inquiry*, **20**, 403 – 417.
- Maturana, H. (1975). The organization of the living: A theory of the living organization. *Intl. J. Man-Machine Studies*, **7**, 313 – 332.
- Maturana, H. (2002). Autopoiesis, structural coupling and cognition: A history of these and other notions in the biology of cognition. *Cybernetics and Human Knowing*, **9**, 5 – 34.
- Maturana, H., & Varela, F. J. (1980). *Autopoiesis and cognition: The realization of the living*. Dordrecht, Holland: D. Reidel Publishing.

- McMullin, B. (2004). Thirty years of computational autopoiesis: A review. *Artificial Life*, **10**, 277 – 296.
- Merleau-Ponty, M. (1962). *Phenomenology of perception*. London: Routledge.
- Merleau-Ponty, M. (1963). *The structure of behaviour*. London: Methuen.
- Moreno, A., & Ruiz-Mirazo, K. (1999). Metabolism and the problem of its universalization. *BioSystems*, **49**, 45 – 61.
- Ono, N., & Ikegami, T. (2000). Self-maintenance and self-reproduction in an abstract cell model. *J. Theoret. Biol.*, **206**, 243 – 253.
- Ruiz-Mirazo, K., & Moreno, A. (2000). Searching for the roots of autonomy: The natural and artificial paradigms revisited. *Communication and Cognition – Artificial Intelligence*, **17**, 209 – 228.
- Ruiz-Mirazo, K., & Moreno, A. (2004). Basic autonomy as a fundamental step in the synthesis of life. *Artificial Life*, **10**, 235 – 260.
- Szostak, J. W., Bartel, D. P., & Luisi, P. L. (2001). Synthesizing life. *Nature*, **409**, 387 – 390.
- Thompson, E. (2004). Life and mind: From autopoiesis to neurophenomenology. A tribute to Francisco Varela. *Phenomenology and the Cognitive Sciences*, **3**, 381 – 398.
- van Gelder, T. (1999). Wooden iron? Husserlian phenomenology meets cognitive science. In Petitot, J., Varela, F. J., Pachoud, B., & Roy, J.-M. (Eds.), *Naturalizing phenomenology*, pp. 245 – 265. Stanford, CA: Stanford University Press.
- Varela, F. J. (1979). *Principles of biological autonomy*. New York: Elsevier, North Holland.
- Varela, F. J. (1991). Organism: A meshwork of selfless selves. In Tauber, A. I. (Ed.), *Organism and the Origin of the Self*, pp. 79 – 107. Netherlands: Kluwer Academic Publishers.
- Varela, F. J. (1996). Neurophenomenology: A methodological remedy for the hard problem. *J. Consc. Studies*, **3**, 330 – 350.
- Varela, F. J. (1997). Patterns of life: Intertwining identity and cognition. *Brain and Cognition*, **34**, 72 – 87.
- Varela, F. J. (1999). The specious present: A neurophenomenology of time consciousness. In Petitot, J., Varela, F. J., Pachoud, B., & Roy, J.-M. (Eds.), *Naturalizing phenomenology*, pp. 266 – 314. Stanford, CA: Stanford University Press.
- Varela, F. J., Maturana, H. R., & Uribe, R. (1974). Autopoiesis: The organization of living systems, its characterization and a model. *BioSystems*, **5**, 187 – 196.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.
- Weber, A., & Varela, F. J. (2002). Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, **1**, 97 – 125.