

From Autopoiesis to Enaction

How a Theory of Life Became a Theory of Experience

Wolfgang Stegemann

Abstract

This paper reconstructs a conceptual shift from the original theory of autopoiesis to later enactive approaches to cognition and experience. The central claim is that the transition is not a straightforward deduction from autopoiesis. Maturana's identification of life with cognition expands the concept of cognition from the outset. Later developments add embodied action, autonomy, adaptivity, agency, sense-making, and first-person experience. Each addition addresses a genuine problem, but the sequence risks turning biological regulation into semantic and eventually phenomenal vocabulary without identifying the mechanism that warrants the transition. The paper therefore distinguishes self-production, regulation, cognition, and experience as different explanatory targets. It argues that autopoiesis remains a powerful account of biological individuality if it is not equated with cognition. An alternative continuation is proposed: instead of locating the decisive development primarily in organism–environment coupling, the analysis follows the evolutionary differentiation of the living system itself, especially the emergence of nervous systems, fast signal integration, recursive closure, and a neurally constituted epistemic operational interior. Embodiment and environmental coupling remain necessary conditions, but they are not thereby mechanisms of experience. The result is a reconstruction of the historical divergence between autopoietic theory and later enactivism and a proposal for a biologically grounded route from living organization to cognition and phenomenal experience.

Keywords: autopoiesis; cognition; enactivism; embodiment; sense-making; adaptivity; experience; consciousness; Maturana; Varela

1. Introduction

The theory of autopoiesis developed by Humberto Maturana and Francisco Varela is one of the most important twentieth-century attempts to define life by organization rather than by a list of properties. A living being is not treated as an aggregate of material parts. It is a self-producing and self-maintaining network in which components participate in the production of the very organization that makes the system an identifiable living unity (Varela, Maturana, & Uribe, 1974; Maturana & Varela, 1980).

This move is theoretically powerful because it changes the explanatory question. Instead of asking which components make something alive, it asks what kind of organization makes those components parts of a living system at all. The organism is not assembled from externally defined functional modules in the manner of a machine. Its components exist within a production nexus that continuously realizes the system itself.

The difficulty begins at another point. Maturana extends the concept of cognition until it coincides with life itself. In *Biology of Cognition* he states that living systems are cognitive systems and that living, as a process, is a process of cognition, explicitly including organisms with and without nervous systems (Maturana, 1970/1980, p. 13). This is a consequential conceptual decision. A form of biological organization is no longer described merely as a precondition of cognition. It is itself called cognition.

In Varela's later work the explanatory focus shifts further. *The Embodied Mind* presents cognition as embodied action and emphasizes recurrent sensorimotor patterns, organism–environment coupling, and the enactment of a world (Varela, Thompson, & Rosch, 1991). Later enactive work adds autonomy, adaptivity, agency, sense-making, normativity, and lived experience. The original question of how a living system constitutes and maintains itself gradually becomes connected to the question of how a meaningful world exists for an organism.

The thesis of this paper is that these are related but distinct explanatory projects. The path from autopoiesis to enaction, sense-making, and experience is not a necessary consequence of the original theory. It consists of conceptual transitions, each of which requires independent justification.

2. The Original Strength of Autopoiesis

The classical formulation of autopoiesis concerns the organization of living systems. An autopoietic system continuously produces components whose interactions regenerate the network that produces them and realizes the system as a unity (Varela, Maturana, & Uribe, 1974).

This establishes a fundamental difference between living organization and technical construction. A machine receives its functional organization from an external context of design. A living organism continually realizes and renews its components within the organization whose persistence is its own persistence.

In this sense the organism possesses an operational interior. This expression does not yet refer to phenomenal experience. It refers to organizational closure. Processes within the system are connected so that their products become conditions for further processes of the same system. The organism remains materially and energetically open, but external events affect it only through the organization that determines how perturbations can be taken up.

Up to this point autopoiesis is a theory of biological organization. No theory of experience follows from it.

3. The First Conceptual Step: Life Becomes Cognition

Maturana goes further. On the same page on which he identifies living with cognition, he characterizes a cognitive system by an organization that defines a domain of interactions in which the system can act in a manner relevant to its maintenance (Maturana, 1970/1980, p. 13).

This changes the extension of cognition dramatically. In ordinary cognitive science the term refers to more specific capacities such as perception, learning, memory, discrimination, problem solving, and thought. Under Maturana's definition, organization-dependent interaction relevant to continued existence is already sufficient.

The objection is not that bacteria or other simple organisms lack complex environmental responsiveness. The issue is what explanatory work is performed by calling such responsiveness cognition. Autopoiesis plus structure-dependent environmental interaction is redescribed as cognition, but no additional mechanism has thereby been identified.

Luisi (2003) made a closely related point in his reassessment of autopoiesis, treating the use of the term cognition for elementary selective coupling as questionable. The biological description may be sound while its cognitive interpretation remains contestable.

4. A Definition Is Not an Explanation

Scientific concepts can of course be stipulated. One may define every self-maintaining, organization-dependent interaction as cognitive. The problem is that a definition cannot substitute for a genetic or mechanistic explanation.

If cognition is definitionally identified with living, the transition from life to cognition disappears from the terminology, not necessarily from biology. The consequences become particularly visible when nervous systems enter the picture. If all life is already cognition, the nervous system can only elaborate an already existing cognitive condition. The possibility that nervous systems introduce a qualitatively new organizational regime is conceptually weakened from the outset.

5. Autopoiesis and Cognition Are Not Identical

Later work has explicitly challenged the identity thesis. Bourguin and Stewart (2004) distinguish autopoiesis from cognition and argue that the two properties need not coincide. Their analysis permits autopoietic systems that are not cognitive and cognitive systems that are not autopoietic.

Bitbol and Luisi (2004) likewise argue, in the context of minimal life, that the lowest level of cognition should not simply be identified with the lowest level of autopoiesis. These critiques are important because they show that the explanatory value of autopoiesis can be retained without accepting Maturana's broad cognitive vocabulary.

At minimum, self-production, regulation, and cognition should be distinguished. Self-production concerns the organizational reproduction of the living system. Regulation concerns processes that keep or move system states within viable ranges. Cognition, if the term is to retain explanatory discrimination, concerns an additional form of organization rather than being a synonym for the first two.

6. An Intermediate Step: Cognition as Effective Action

The development toward enactivism was gradual. In *The Tree of Knowledge* Maturana and Varela characterize knowing in terms of effective action and connect cognition with the bringing forth of a world (Maturana & Varela, 1987).

This already shifts the explanatory emphasis. Cognition is not treated primarily as internal representation. It is relational, enacted in the history of interactions through which an organism continues to exist. The move remains compatible with structural coupling, but the question has changed. It is no longer only how the system maintains itself, but how a world becomes relevant within its activity.

7. The Transition to Enaction

In *The Embodied Mind* the shift becomes explicit. Varela, Thompson, and Rosch (1991) develop cognition as embodied action. Perception and action are mutually dependent, and cognitive structures emerge from recurrent sensorimotor patterns rather than from the passive recovery of a pregiven external world.

This is a persuasive challenge to naive representationalism. Organisms do not receive ready-made descriptions of their environments. Perturbations alter sensory structures, those alterations are processed according to the organization of the system, and action changes the conditions of subsequent perturbation.

Yet the rejection of naive representation is not itself a positive mechanism of cognition or experience. Showing that cognition is situated in organism–environment dynamics does not establish what makes those dynamics cognitive in the stronger sense, and still less what makes them experienced.

8. Where the Break Lies

The original autopoietic question can be formulated as: how does a living system constitute and maintain itself as a unity? The enactive question increasingly becomes: how does a world of relevance arise for a living system through its activity?

The questions are connected, but they are not identical. Villalobos and Palacios (2021) have argued that autopoietic theory and enactivism should not simply be treated as possessing the same mark of the cognitive. Their analysis supports the claim that the later program cannot be read as a mere terminological continuation of the earlier one.

The word break should therefore be used systematically rather than biographically. Varela did not need to announce a repudiation of Maturana for a shift of explanatory target to occur. The break consists in the movement from self-constitution toward embodiment, sensorimotor world-involvement, meaning, and experience.

9. From the System to a Meaningful World

Once this shift is recognized, an alternative historical possibility becomes visible. Starting from autopoiesis, one could have investigated the increasing internal differentiation of living systems: specialized signaling, nervous systems, integration of sensory and motor domains, recurrent neural activity, and the loss of causal independence among previously segregated functions.

Enactivism instead places increasing explanatory weight on the relation between organism and environment. The trajectory can be schematically represented as autopoiesis, autonomy, enaction, and sense-making. With each step the vocabulary becomes more semantic.

10. From Autopoiesis to Normativity

Weber and Varela (2002) develop the connection between autopoietic individuality, natural purpose, and biological normativity. Conditions are not neutral for a living system. Relative to its continued organization, some states are favorable and others destructive.

This is biologically important, but it requires a distinction between two senses of 'for the system'. A condition can be functionally relevant for the system, and a condition can appear to the system as something. The first is a biological claim. The second is phenomenal. Functional relevance does not by itself entail appearance.

11. The Problem of Sense-Making

Sense-making becomes a central enactive term for the organism-dependent significance of environmental conditions. An organism does not merely undergo physical impacts. Because it has its own viability conditions, different events matter differently to it.

The ambiguity lies in the word meaning. If sense-making means that environmental conditions have different functional consequences for viability, the concept remains biological. If it means that something appears as significant to the organism, phenomenal content has already entered the description.

The inference from 'X affects the viability of S' to 'X is experienced by S' is invalid without an additional account. The first proposition is available from the observer's biological description. The second attributes experience.

12. Di Paolo's Adaptivity and the Acknowledged Gap

Di Paolo (2005) is particularly revealing because he explicitly argues that the original formulation of autopoiesis must be developed further if it is to account for sense-making. His concept of adaptivity describes the capacity of a system to regulate its relation to its own viability boundaries.

This is theoretically significant. If autopoiesis already yielded cognition and meaning, adaptivity would not be needed as an additional explanatory category. Its introduction indicates that the transition from self-production to sense-making remains open.

Adaptivity can explain a richer form of regulation. It does not yet explain experience. Regulation of viability is not the same as experiencing viability.

13. The Gradual Semantization of Biological Processes

A general pattern can now be identified. An environmental condition causes a biological change. It is then described as relevant to the organism. Relevance becomes value, value becomes meaning, meaning becomes sense-making, and sense-making is placed in the vicinity of lived experience.

Each step sounds plausible because, in human beings, functional relevance, meaning, and experience often coincide. Metabolic deficit is functionally relevant, changes behavior, and can be experienced as hunger. But it does not follow that every instance of metabolic relevance already contains an elementary form of hunger or experience.

An evolutionary precondition is not identical with the phenomenon that later develops from it.

14. Life Is a Condition of Cognition, Not Cognition Itself

Autopoiesis can be interpreted differently without diminishing its importance. It can be treated as the foundational organization of living systems rather than as cognition itself.

On this view the sequence is developmental rather than definitional: autopoiesis gives rise to a living organism; living organization supports regulation; evolution produces specialized signaling and nervous systems; nervous systems make possible increasingly integrated neural organization; cognition and, under further conditions, experience emerge from these new organizational regimes.

This model preserves continuity without identity. Life is the organizational basis within which cognition can arise, but life is not thereby cognition. Cognition in turn is not automatically consciousness.

15. An Alternative Route: From a Chemical to an Epistemic Operational Interior

Autopoiesis establishes a chemically operational interior. The term interior refers to the organization by which the living system produces and maintains itself, not to subjective space.

With nervous systems a new form of organization arises within that living system. Perturbations of different origins can be processed rapidly, related to one another, and distributed through specialized networks. Sensory, motor, interoceptive, and regulatory processes become neurally mediated.

This creates what may be called an epistemic operational interior. Epistemic does not mean that the organism possesses an objective representation of an external world. It means that external and internal perturbations are transformed into system-specific neural patterns and become available to further processes only in that transformed form.

The explanatory question now changes from 'How does interaction generate meaning?' to 'Which changes in the internal organization of the living system enable the transition from regulation to cognition and eventually to phenomenal organization?'

16. Embodiment Is a Condition, Not an Explanation

Neural processes are unquestionably embodied. Nervous systems develop inside organisms and depend on metabolism, endocrine regulation, immune processes, sensory input, and movement. Bodily action alters subsequent sensory conditions.

These facts establish causal embedding. They do not yet identify the mechanism of experience. A necessary condition is not identical with the mechanism of the phenomenon for which it is necessary.

The body supplies conditions, perturbations, regulatory states, and possibilities of action. The question of what organizational state is experienced remains. Embodiment is therefore indispensable to an account of cognition without automatically constituting an explanation of consciousness.

17. Neurophenomenology and the Move to Experience

In Varela's later neurophenomenology, first-person experience becomes an explicit methodological concern. Varela (1996) treats first-person experience as an irreducible phenomenal domain and proposes reciprocal constraints between disciplined phenomenological reports and cognitive neuroscience.

This is a legitimate methodological program. Phenomenological reports can guide experiments and can be correlated with neural states. Yet correlation and mutual constraint do not provide a genesis mechanism. The relation between descriptions does not by itself explain why a particular neural organization is experienced.

The distance from the initial autopoietic problem is therefore striking. The theory begins with the organization of the living system and arrives at a methodology for relating neural and experiential descriptions. The mechanism connecting the two remains to be specified.

18. The Blind Spot of Enactivism

The blind spot of enactivism is not its attention to body, action, or environment. Those are strengths. The problem is the temptation to treat the conceptual sequence from autonomy through adaptivity, agency, and sense-making as if it already explained the transition from biological organization to experience.

The arrows in that sequence do not all have the same status. Some mark biological dependencies, some theoretical extensions, and some transitions between levels of description. The final transition remains especially open.

A system may avoid a state because that state threatens its viability. This can explain behavior. It does not explain why danger feels like anything. An organism may seek food because metabolic conditions alter regulatory activity. This can explain behavior. It does not explain hunger as experience.

19. The Problem of Levels of Description

The deeper methodological issue is the mixing of explanatory levels. Autopoiesis is an organizational-biological concept. Adaptivity is regulatory. Sensorimotor coupling describes dynamics between perception and action. Sense-making introduces semantic vocabulary. Lived experience is phenomenological.

Relations among these descriptions are legitimate, but a relation between descriptions is not automatically a mechanism. A phenomenal concept is not biologically explained merely because a biological process is redescribed in phenomenal or semantic language.

Calling metabolic relevance meaning does not explain meaning. Calling adaptive regulation sense-making does not explain experience. Calling sensorimotor coupling embodied cognition still leaves open what makes the relevant process cognitive.

20. Autopoiesis and the Extended Mind

The extended mind thesis itself was formulated by Clark and Chalmers (1998), not by Varela. It should therefore not be presented as a direct component of Varela's theory. There is nevertheless a conceptual affinity between some enactive approaches and later externalist accounts.

If cognition is defined primarily through dynamic organism–environment relations, it becomes natural to ask whether environmental components can themselves become parts of a cognitive system. Yet this produces a tension with the original autopoietic emphasis on the constitution of a system as a unity.

Coupling does not abolish the distinction between system and environment. Causal participation is not automatically constitutive membership. From a strictly autopoietic perspective, the expansion of a cognitive system beyond the living system therefore requires additional argument.

21. Another Continuation of Autopoiesis

An alternative program would continue autopoiesis by examining evolutionary transformations within autopoietic systems. Nervous systems introduce new speeds and ranges of causal coupling. Sensory, motor, interoceptive, and regulatory processes can become integrated. Increasing centralization allows previously segregated processes to converge, while recurrent connectivity permits neural states to modify one another continuously.

The decisive question then concerns the organizational conditions under which formerly separable functional domains lose causal independence and form a non-factorizable global state. A theory of experience would begin here.

The central mechanism would not be environmental coupling as such, embodiment as such, or meaning as a theoretical label. It would be a specific organization within a living neural system. The original insight of autopoiesis would thereby be retained and biologically extended.

22. Autopoiesis Without Mental Inflation

Separating life, cognition, and experience also prevents conceptual inflation. If life is cognition and cognition is subsequently characterized through sense-making, intentionality, and experience, mental vocabulary can be projected downward onto minimal forms of life.

A bacterium lives. It is autopoietically organized. It responds selectively to its environment. It regulates internal states and may display sophisticated adaptive behavior. None of these propositions entails that it experiences.

This does not diminish life. It allows biological organization to be investigated without loading it with concepts whose primary explanatory role belongs elsewhere.

23. The Historical Point

The historical development can now be read in two ways. One can regard enactivism as the progressive enrichment of a unified autopoietic research program. But one can also see the later theory as reintroducing distinctions that the early equation of life and cognition had initially flattened.

After life is identified with cognition, cognition is enriched through effective action, embodiment, enaction, normativity, adaptivity, sense-making, and finally neurophenomenological engagement with experience. Di Paolo's explicit need to supplement autopoiesis with adaptivity is especially revealing: the original identity claim had not supplied the explanatory resources required by the later theory.

Had life and cognition remained distinct from the beginning, a different research program would have remained open: autopoiesis for living organization, nervous-system evolution for cognitive organization, and a further theory of neural integration for phenomenal experience.

24. Conclusion

Autopoiesis should not be rejected together with every later enactive extension. Its original insight remains powerful: a living organism is a self-producing and self-maintaining process nexus whose components can be understood adequately only within that organization.

The problematic step begins when this organization is itself called cognition. From there a conceptual chain develops from life to cognition, embodied action, sense-making, and experience. None of these transitions is self-evident.

The alternative is not to remove body and environment from the account. The body enables and regulates neural systems. The environment perturbs the organism and structures its development. Action changes the conditions of subsequent perception. All of this is indispensable.

But conditions of a phenomenon are not identical with its mechanism. The decisive continuation of autopoiesis may therefore lie less in extending cognition outward into body and environment than in reconstructing the evolutionary differentiation of the autopoietic system itself.

The chemically operational interior of life is supplemented, with nervous systems, by an epistemic operational interior. Only at that level does the question arise which organization of neural processes enables the transition from regulation to cognition and from cognition to experience.

Varela's later program increasingly investigates embodiment, action, world-involvement, and the methodological integration of experience. An alternative program locates the decisive transition in the organization of the living system itself. At this fork lies the difference between describing the conditions and relations of experience and explaining its biological genesis.

References

- Bitbol, M., & Luisi, P. L. (2004). Autopoiesis with or without cognition: Defining life at its edge. *Journal of the Royal Society Interface*, 1, 99–107. <https://doi.org/10.1098/rsif.2004.0012>
- Bourgine, P., & Stewart, J. (2004). Autopoiesis and cognition. *Artificial Life*, 10(3), 327–345. <https://doi.org/10.1162/1064546041255557>
- Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7–19. <https://doi.org/10.1093/analys/58.1.7>
- Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 4, 429–452. <https://doi.org/10.1007/s11097-005-9002-y>

- 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: Toward a New Paradigm for Cognitive Science*. MIT Press.
- Luisi, P. L. (2003). Autopoiesis: A review and a reappraisal. *Naturwissenschaften*, 90(2), 49–59. <https://doi.org/10.1007/s00114-002-0389-9>
- Maturana, H. R. (1970/1980). Biology of cognition. In H. R. Maturana & F. J. Varela, *Autopoiesis and Cognition: The Realization of the Living*. Dordrecht: D. Reidel.
- Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and Cognition: The Realization of the Living*. Dordrecht: D. Reidel.
- Maturana, H. R., & Varela, F. J. (1987). *The Tree of Knowledge: The Biological Roots of Human Understanding*. Boston: Shambhala.
- Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: Harvard University Press.
- Varela, F. J. (1996). Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies*, 3(4), 330–349.
- Varela, F. J., Maturana, H. R., & Uribe, R. (1974). Autopoiesis: The organization of living systems, its characterization and a model. *BioSystems*, 5(4), 187–196. [https://doi.org/10.1016/0303-2647\(74\)90031-8](https://doi.org/10.1016/0303-2647(74)90031-8)
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
- Villalobos, M., & Palacios, S. (2021). Autopoietic theory, enactivism, and their incommensurable marks of the cognitive. *Synthese*, 198, S71–S87. <https://doi.org/10.1007/s11229-019-02376-6>
- 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. <https://doi.org/10.1023/A:1020368120174>