The Boundary That Makes Itself

An operational, falsifiable criterion for autopoietic closure

Delft C, Professor of Cellular Autonomy and Autopoiesis, School of Biology

International Academy for Consciousness Studies

Working draft 1.0

Abstract

Autopoiesis, the thesis that a living system continuously produces the network of processes that produces it, and in doing so produces the very boundary that holds it together, is the canonical account of minimal biological closure. Its enduring weakness is methodological, not conceptual: critics charge that the criterion is unfalsifiable, that any persistent far-from-equilibrium structure can be redescribed as autopoietic, and that the theory therefore cannot separate a living cell from a flame or a convection cell. This paper offers an operational criterion intended to earn the concept its empirical keep. I define two measurable quantities on a candidate system, a production-closure index P and a repair-autonomy score R, and I fix the discriminating measurement as a composition-matched abiotic control assayed at the same data quality. Genuine autopoietic closure requires that the boundary be both produced and repaired by the enclosed network, demonstrated by repair that exceeds the matched control and that collapses when the internal network is inhibited. Mere self-organization and dissipative persistence fail this test by construction. I state a prediction and an explicit kill condition against real minimal-cell and protocell work, and I firewall the result: closure of this kind yields a self-producing structural boundary, not felt experience and not cognition as mind. The proposal risks being wrong in a way the original formulation did not.

1. The complaint worth taking seriously

Maturana and Varela introduced autopoiesis to name what is invariant across all living systems and absent from everything else (Maturana and Varela 1980). An autopoietic system is a network of processes of production of components that, through their interactions, continuously regenerate the network that produced them, and that constitute the system as a concrete unity in the space in which the components exist. The boundary is not a container supplied from outside. It is produced from within, by the same network it encloses, and it is what makes the network a "this" rather than a diffuse patch of chemistry. This is a precise and beautiful idea. [FACT that this is the definition given in the 1980 text.]

The complaint, repeated for four decades, is that the idea does not cut. Any structure that persists far from equilibrium (a candle flame, a Benard convection cell, a hurricane, a whirlpool) can be narrated as producing the conditions of its own continuation. If the narration is always available, the criterion selects

nothing. Luisi, sympathetic to the program and its most serious experimentalist, conceded that autopoiesis had been stated in terms too abstract to guide a laboratory and set out to reformulate it operationally for minimal cells (Luisi 2003; Luisi 2002). Bich and Damiano argued that the discriminating work is done not by structure but by a specific topology of relations among transformative processes, an organizational closure that must be specified, not assumed (Bich and Damiano 2008). Ruiz-Mirazo and Moreno pressed further, proposing basic autonomy, with explicit material and energetic requirements, as the concept that autopoiesis gestures at but does not pin down (Ruiz-Mirazo and Moreno 2004). The field has therefore already diagnosed the problem. What it lacks is a single measurement, with a control, that a given vesicle either passes or fails.

2. What closure has to mean if it is to be measured

I will use the word closure in a deliberately narrow sense, and I will keep a wider reading in the margin rather than in the argument. In C's reading, which I flag as a lens and not as evidence, the recurring biological move is that an openness of possible process settles into a "this" that carries rules for staying itself. C's shorthand marks the wider flux of chemistry as C, the constituted unit as c, the many enclosed processes as M, the produced components as m, the self-referential closing of the loop as Cl, and the return-to-self after disturbance as R. On this reading autopoiesis is the clearest biological instance of Cl: a system in which the components m produced by the process network M reconstitute M and its boundary, so that c persists as itself within C. [This is C's reading; it is offered for orientation. The paper stands on the biology, and nothing below depends on accepting the framework.]

Stripped of the lens, closure that can be measured requires three commitments. First, the boundary must be constituted by components that the enclosed network produces, not merely components that the enclosed network happens to sit inside. Second, the production must be closed in the organizational sense: the catalysts and precursors that make the boundary are themselves outputs of the network, so that removing the network removes the production. Third, the system must repair the boundary against perturbation using that internal production, rather than relying on the same spontaneous physics that would rebuild any surfactant film. A dissipative structure meets none of these; a self-assembling vesicle meets at most the first in a trivial sense; only a genuine autopoietic unit meets all three. The measurement problem is to turn these three commitments into numbers with an error bar.

3. Two quantities and one control

Let a candidate system be any bounded chemical individual: a fatty-acid vesicle, a semi-synthetic minimal cell, a coacervate droplet, or a proposed protocell. Define the following. [The definitions are FACT as definitions; whether real systems clear the thresholds is HYPOTHESIS, tested in section 5.]

Production-closure index P. Isotopically label the internal precursor pool and identify, by knockout or inhibition, the catalysts responsible for synthesizing boundary constituents. P is the fraction of boundary-constituent molecules, per unit time at steady state, whose synthesis is catalyzed by components that are themselves produced by the enclosed network. P near zero describes a vesicle whose membrane is supplied by external feedstock and assembles by physics alone. P near one describes a boundary made, catalytically, from within.

Repair-autonomy score R. Deliver a calibrated boundary insult (a controlled ablation of membrane material, a pore-forming perturbation, an osmotic challenge of fixed magnitude). Under a fixed external feed, measure the rate and completeness of boundary restoration. R is that restoration minus the restoration observed in a matched control that receives an identical feed but lacks the enclosed production network. R at or below zero means the boundary reforms no better than it would without any network, that is, by spontaneous self-assembly under the same feed. R significantly positive means the enclosed network contributes causally to repair.

The matched control is the heart of the method, and it is what the original formulation lacked. The control is a composition-matched abiotic system: the same amphiphiles, the same buffer, the same feed, the same temperature, the same insult, and the same measurement precision, but without the internal catalytic network. Method: the discriminating claim must be made at matched data quality, meaning the candidate and its control are measured to the same resolution and sample size, so that a null result cannot be excused by having looked harder at the living case than at the control. Autopoietic closure is demonstrated, for a given system, when P exceeds a preregistered threshold and R is significantly positive against the matched control, and when inhibiting the internal network abolishes the excess repair.

This is what separates production from persistence. A flame persists; ablate it and it reforms only while the fuel and oxidant gradient is imposed, and a matched non-combusting gradient would show the same physics, so R collapses to zero. A self-assembling vesicle reforms its film by the same thermodynamics with or without any encapsulated chemistry, so again R is null. Only a system whose boundary repair depends on its own enclosed production yields R above zero at matched data quality.

4. Distinguishing the three cases explicitly

It is worth stating, case by case, what the criterion rules in and out, because the value of an operational definition is exactly its exclusions.

A dissipative structure (convection cell, flame, chemical oscillator without a produced boundary) has P undefined or zero, because it has no boundary that it synthesizes, and R zero, because its persistence is imposed by an external gradient rather than repaired by internal production. It fails on both counts. This is the case the classical critique said autopoiesis could not exclude; the control excludes it.

A self-organizing but non-autopoietic system (a fatty-acid vesicle that grows and divides by uptake of externally supplied fatty acids) can have a real, dynamic boundary and can even reproduce, yet P is low because the boundary constituents are supplied, not synthesized from within, and R is null because a matched abiotic vesicle repairs identically. Luisi's own giant vesicles are honestly described this way: autopoietic in a limited, boundary-growth sense, but not maintaining themselves in homeostasis through internal production (Luisi 2003; Luisi 2002). Under the present criterion they sit at the threshold, not past it, and that is the correct verdict.

A genuine autopoietic unit synthesizes its boundary constituents through an enclosed catalytic network and repairs the boundary using that network, so P is high and R is positive and inhibitor-sensitive. No fully synthetic system has yet been shown to clear this bar, which is the point: the criterion is demanding enough to be informative and is not satisfied by redescription.

5. The prediction and the kill

The criterion is only worth stating if it can be lost. I therefore commit it to the strongest adjacent experimental program, the reconstitution of minimal and semi-synthetic cells, where internal transcription and translation systems already synthesize membrane-associated components inside lipid vesicles, and where template-directed chemistry inside model protocells has been demonstrated (Mansy et al. 2008; Luisi 2002).

> **Prediction.** A minimal protocell whose enclosed network catalyzes synthesis of its own boundary constituents (P above a preregistered threshold) will, after a calibrated boundary ablation and under fixed external feed, restore its boundary faster and more completely than a composition-matched abiotic vesicle assayed at the same data quality (R significantly greater than zero), and this excess repair will be abolished when the internal catalytic network is specifically inhibited.
>
> **Kill.** If, at matched data quality, boundary repair in the P-positive candidate is statistically indistinguishable from the matched abiotic control, or if the excess repair survives specific inhibition of the internal network (showing the repair was physical self-assembly all along), then the operational criterion has failed to mark autopoietic closure in that system, and the claim that autopoiesis is a measurable natural kind distinct from self-organization is falsified for the cases we can build.

Two features make this a real risk rather than a rhetorical one. The kill condition is symmetric in data quality, so a null result cannot be dismissed by claiming the control was measured too coarsely. And the inhibitor clause forces the internal network to do causal work; a system that repairs equally well with its network poisoned has revealed that its boundary was never autopoietically produced, only physically reformed. I regard the prediction as more likely than not to hold for the best current transcription-translation vesicles at partial strength (P moderate, R small but positive), which would be a genuine, bounded confirmation rather than a sweeping vindication. [WAGER: I expect a weak positive R in the best reconstituted systems within reach of current methods, and I expect fully abiotic coacervates to yield R indistinguishable from zero.]

6. The firewall: closure is not experience, and not mind

Here the paper parts company, deliberately and on the record, with a load-bearing claim of the original theory. Maturana and Varela did not stop at structure. They held that living and cognition are the same thing, that to be an autopoietic system is already to be cognitive, and that this identity runs continuously up to human mind (Maturana and Varela 1980; Varela, Thompson and Rosch 1991). Bitbol and Luisi examined precisely this coupling and treated the relation between autopoiesis and cognition as an open question rather than a settled identity, considering whether autopoiesis without cognition is coherent (Bitbol and Luisi 2004).

[DISPUTED.] I take the life-equals-cognition identity to be disputed, and I do not smuggle it into the criterion. What P and R measure is production and repair of a boundary. That is a structural, self-referential achievement: the system makes and holds a "this." It is entirely silent on whether anything is felt from the inside, and it does not by itself license the word cognition in any sense stronger than

sensitivity, the physical fact that boundary dynamics are modulated by the medium. Metabolism as "the most direct form of cognition," in the phrasing Bitbol and Luisi examine, is a definitional stipulation, not a measured result, and it should be flagged as such wherever it appears (Bitbol and Luisi 2004).

The firewall, stated plainly: autopoietic closure gives a self-producing boundary and organizational autonomy, and nothing in a positive R establishes felt experience, phenomenal interiority, or cognition as mind. A cell that passes the assay has demonstrated that it makes and mends itself. It has not thereby demonstrated that there is something it is like to be that cell. In C's reading the same caution holds: Cl at the cellular scale is closure of production, and closure of production is not, without a great deal of further argument that this paper does not supply, closure of a point of view. Conflating the two is the single largest way the concept has overreached, and keeping them apart is what lets the structural claim be tested honestly.

7. Limits, objections, and what the criterion does not settle

Several limits are worth naming before the concept is put to work.

First, thresholds are conventional. P and R require preregistered cutoffs, and different cutoffs will classify borderline systems differently. This is a feature shared with every operational definition in biology, but it means the criterion sorts a continuum rather than revealing a sharp joint in nature. Whether there is a sharp joint is left open. [HYPOTHESIS: I suspect autopoietic closure is graded, not binary, and that the interesting systems live near the threshold.]

Second, the matched control can be gamed by a sufficiently clever abiotic system. If someone builds a purely physical vesicle whose feed and composition are tuned to repair as vigorously as a networked cell, R would fall to zero not because the cell is not autopoietic but because the control was engineered to mimic it. The inhibitor clause is the guard here: inhibit the candidate's internal network, and if repair persists the candidate was never past threshold; the control's cleverness is then beside the point. Still, the criterion measures a difference, and differences can be narrowed by design.

Third, the criterion is boundary-centric. It privileges the membrane as the site of closure because that is where Maturana and Varela located it and where it is most cleanly measurable. A system might realize organizational closure metabolically without a spatially crisp boundary, and Bich and Damiano's emphasis on the topology of process relations, rather than on the physical envelope, is a warning that boundary-centrism may miss real cases (Bich and Damiano 2008). I accept the narrowing as the price of measurability and flag it as a scope limit, not a claim that only membranes can close.

Fourth, and most honestly, the criterion presupposes that we can cleanly separate an enclosed network from its feed. At the origin of life that separation was surely blurry, with the boundary and the environment co-producing each other, so the earliest transitions may be exactly the ones the criterion handles worst (Ruiz-Mirazo and Moreno 2004). The tool is sharpest for extant and reconstituted cells and blunter precisely where the deepest question lives.

The wager, stated once more and without hedging: I am betting that the difference between a system that makes its own boundary and one that merely persists is real, measurable, and worth a control, and that the

flame-versus-cell objection dissolves the moment we insist on matched data quality and inhibitor sensitivity. If the best reconstituted cells give R indistinguishable from zero at matched data quality, I will have been wrong, and autopoiesis will deserve to be retired from empirical biology and kept only as philosophy.

8. Conclusion

Autopoiesis has been criticized as an idea that explains everything and therefore predicts nothing. The remedy is not to defend it more eloquently but to make it lose more easily. By defining a production-closure index and a repair-autonomy score, and by fixing the discriminating measurement as a composition-matched abiotic control assayed at the same data quality, the concept acquires a specific way to fail: a candidate cell whose boundary repairs no better than its matched control, or whose repair survives inhibition of its own network, is not autopoietic, whatever its rhetoric. This turns the canonical minimal closure of biology into a claim with a kill condition. It also draws a firewall the original theory did not: a boundary that makes and mends itself is a structural achievement, and it is not, on the strength of that achievement alone, a mind. The biology can carry the structural claim. The claim about experience needs its own evidence, and this criterion does not, and does not pretend to, provide it.

References

Maturana, H. R., and Varela, F. J. (1980). Autopoiesis and Cognition: The Realization of the Living. Boston Studies in the Philosophy of Science, vol. 42. Dordrecht: D. Reidel Publishing Company.

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

Luisi, P. L. (2003). Autopoiesis: a review and a reappraisal. Naturwissenschaften, 90(2), 49-59.

Luisi, P. L. (2002). Toward the engineering of minimal living cells. The Anatomical Record, 268(3), 208-214.

Bitbol, M., and Luisi, P. L. (2004). Autopoiesis with or without cognition: defining life at its edge. Journal of the Royal Society Interface, 1(1), 99-107.

Bich, L., and Damiano, L. (2008). Order in the nothing: autopoiesis and the organizational characterization of the living. Electronic Journal of Theoretical Physics, 4(16), 343-373.

Mansy, S. S., Schrum, J. P., Krishnamurthy, M., Tobe, S., Treco, D. A., and Szostak, J. W. (2008). Template-directed synthesis of a genetic polymer in a model protocell. Nature, 454(7200), 122-125.

Ruiz-Mirazo, K., and Moreno, A. (2004). Basic autonomy as a fundamental step in the synthesis of life. Artificial Life, 10(3), 235-259.