Exo-autopoietic bodies: The quest for the theoretical identity of living beings

1. Introduction

Despite all the encyclopedic knowledge that biological sciences have accumulated regarding living beings, their physiology and behaviour, their molecular bases, their development and evolution, it is still frustratingly elusive to find a neat and uncontroversial answer to the (apparently) simple question "What are living beings?" The traditional approach to answering this question has been by means of definitions. Many have been proposed in the literature over the years (each one emphasising different aspects of living beings, such as biochemical composition, metabolism, thermodynamics, evolution, or self-organisation), but none have achieved transversal acceptance in the community (Sagan 1970; Pályi, Zucchi and Caglioti 2002; Tsokolov 2009; Bedau and Cleland 2010; Trifonov 2011; Kolb 2018). So much is the case that some have declared, with resignation, that it is impossible to find such a definition and that we should better forget the whole question (Machery 2012).

Why this systematic failure in answering the question of living beings? One diagnosis is that the problem lies not in the theoretical content of the many definitions but in their common form, to wit, that they are all definitions. According to Cleland and Chyba (2002; 2007), definitions are linguistic/conceptual tools designed to deal with the meaning of words, not directly with the entities of the natural world, so they fail to fix natural kinds (Cleland and Chyba 2002; 2007). Under the assumption that living beings constitute a natural kind, the conclusion is that searching for a definition of life is the wrong approach. The alternative would be to follow the example of sciences such as physics and chemistry, which, instead of building definitions for the natural phenomena and kinds they study, state theoretical identities (Cleland and Chyba 2007; Cleland 2019b).

This diagnosis suggests that the question of living beings should be faced not with a definition but with a theoretical identity statement.[1] However, although promising from the formal point of view, this move would face other problems. According to Cleland and Chyba (2007; Cleland 2019b), theoretical identities are possible and pertinent only against the background of a general scientific theory, which, in turn, requires a robust empirical basis. Current biology, they argue, does not have enough empirical basis to formulate a general theory of life and, therefore, cannot validly formulate a theoretical identity for living beings. In their view, the underdeveloped state of current biological science makes any attempt to fix a theoretical identity for living beings pointless or inconveniently premature (Cleland 2019a, b).

In this paper, contrary to this latter recommendation, I make room for and propose a theoretical identity statement for living beings. First, I argue that the limitations and risks that Cleland and Chyba see in proposing such an identity are not as severe as they assume. Second, I suggest answering the question, "What are living beings?" with the theoretical identity statement, 'Living

---

[1] There is also the possibility of defending definitions and insisting on their utility (see, e.g. Mix 2015; Bich and Green 2018). I will not follow that strategy in this article.

beings are exo-autopoietic bodies.' I explain this proposal's theoretical grounds and defend it against potential objections.

2. The possibility of a theoretical identity for living beings

Under the assumption that living beings constitute a natural kind,[2] Cleland and Chyba argue that the definitional approach to the question of living beings fails because definitions, as a general rule, are not suited to fix natural kinds (Cleland and Chyba 2002, 2007; Cleland 2006). Definitions, they explain, are concerned primarily with concepts and language, not directly with the entities that populate the natural world (Cleland 2006). Definitions, because of that, may work well to fix non-natural (i.e., conventional) kinds, for which we can stipulate neat and unambiguous criteria. However, in the case of natural entities, the essential properties of which we come to know not by stipulation but by means of long and hard scientific research, definitions are not the best tools. Because they typically aim to establish necessary and sufficient conditions to fix a category, they remain open to ambiguities, borderline cases and counterexamples. Moreover, Cleland (2019a, b) adds that definitions are theoretically poor as they only serve classificatory purposes in circumstances where, in trying to understand the nature of something, science aims to explain it rather than merely classify it. The alternative? Looking at the history and development of well-established sciences (i.e., physics, chemistry), Cleland and Chyba (2002, 2007) observe that natural kinds are fixed by means of theoretical identities. The paradigmatic case is chemistry's discovery and statement that water is H_2O. Instead of proposing a definition of water, chemistry answers the question "What is water?" with the theoretical identity statement 'Water is H_2O.' Thus, following this and similar examples in sciences, we are invited to explore and find not a definition but a theoretical identity for living beings.

However, how we read the invitation is critical. Cleland (2019a, b) gets deep into the history of sciences and extracts what she takes as a canonical form of scientific theoretical identity, i.e., a model we should replicate in every respect if we are to elaborate and validate a theoretical identity for living beings. I, instead, think that the invitation is not to replicate the model of other sciences in every respect but to adapt it according to the specific characteristics of the question of living beings.

Theoretical identities, according to Cleland (2019b), (i) correspond to empirical discoveries, (ii) are accepted based on the explanatory and predictive, not merely classificatory, power they offer, and (iii) emerge in the context of a well-established, general theory. Let us review these points.

Although criteria (i) and (ii) may well be the case in the history of sciences such as physics and chemistry, it does not follow, absent additional arguments (which Cleland does not provide), that proposals of theoretical identity in other sciences need to meet them forcefully. It is critical how we interpret the lessons from the history of sciences. It is one thing to take the example of other sciences as a general model; it is another to assume that they must be imitated in every respect.

---

[2] This assumption, which I will follow in this article, is not without contestation. See, for instance, Hermida (2016), and Mariscal and Doolittle (2020).

When we formulate the question “What are living beings?” we ask something similar but also something different to what chemistry, in its moment, asked with the question “What is water?” We ask something similar because we want to find -if we follow Cleland and Chyba’s advice (which I think we should follow)-, the theoretical identity, not a definition, of life. Yet, we ask something different because each question's epistemic demands and contexts are different. The chemistry of the 1700s and 1800s needed to know what water was made of. It already knew several of its observable properties but ignored the deep (micro) structure that could explain those properties. In that context, the discovery that water was made of hydrogen and oxygen atoms came as an empirical finding that contributed to explaining water’s observable properties. The epistemic gap was empirical and explanatory, and the answer to the question “What is water?” was expected to fill that gap.

The case with living beings is different. Current biology already knows what living beings are made of, down to the chemical details, and can explain, based on that and other pieces of knowledge, most of the observable phenomena associated with life. If the question “What are living beings?” is still with us, it is not because we are still pursuing an empirical discovery about their microstructure, some unknown material, energy or force that will explain their properties. The question’s motivation, I submit, is more about understanding and clarifying the place of living beings in nature, that is, their distinctiveness as a natural kind. It is, if you will, more a classificatory enterprise than an explanatory one. And there is nothing necessarily wrong with that because it corresponds to how biology has developed its knowledge regarding living beings.

If a theoretical identity, within a determined historical, scientific context, comes as an empirical discovery that explains phenomena that we did not know how to explain, that is great, but that does not mean all theoretical identities in all sciences must do the same. Certainly, proposing a theoretical identity for a natural kind must be done in accordance with the scientific empirical knowledge accumulated regarding said kind, but this is different from asking that the theoretical identity be itself the product of empirical discovery. Similarly, the proposal of a theoretical identity must be in accordance (unless it comes with new empirical evidence and explanatory power) with the accepted explanatory apparatus of the corresponding scientific discipline. Yet, this is different from asking that the theoretical identity explains something (until then) unexplained.

In general, the proposal of a theoretical identity for living beings should be reviewed according to how it works for the purposes and epistemic demands of its own scientific context, not according to how it fits with an alleged ideal model of theoretical identities in the history of science. Cleland may be too focused on the canonical form of how other, more consolidated sciences than biology have established their theoretical identities. I propose to be more pragmatic. If a proposal of theoretical identity for living beings is grounded on respectable scientific knowledge, helps us to understand the distinctiveness of living beings as a natural kind, provides an adequate level of generality, and resists counterexamples, then the point of whether or not it meets the (presumably) ideal characteristics of theoretical identities in the history of science should be irrelevant.

Regarding criterion (iii), Cleland and Chyba (2007; Cleland 2019b) argue that theoretical identities are possible and pertinent only in mature enough sciences, meaning sciences that have already developed a universal theory. Premature proposals of theoretical identity, they warn, risk embracing wrong, premature ontologies, which can seriously compromise the development of science.

According to the authors, biology lacks a well-established, universal theory of life because its empirical base is not (yet) diverse and big enough. All the samples of life with which biology has built its theorisation are strictly local (earthly), which, arguably, represent a tiny part of the potential diversity of life in the universe. This is what Cleland calls the “N = 1 problem” (2019b). That is the problem of having only one sample of the phenomenon under study. Biology, Cleland goes on, lacking a universal theory of life, is not in a position to propose (let alone validate) a theoretical identity for living beings. Trying to force such a theoretical identity at this premature stage would hinder rather than promote the development of biological science. Instead, we should try to access more diverse (non-earthly) samples of life so we can have a more representative empirical base to build a universal theory of living beings. This means, advises Cleland (2019b), we should promote astrobiological explorations. In the absence of a general theory of life, she proposes to explore space in an open-minded way, equipped with tentative criteria to detect biosignatures of unexpected or anomalous forms of life (Cleland 2019a, 2019b).

The idea of exploring space with tentative criteria of biosignatures sounds reasonable, and I have no objections to it. My point is that this initiative should be developed jointly with, not instead of, the proposal and discussion of theoretical identities for living beings. These latter are not as inviable and inconvenient as Cleland and Chyba think they are. If well-constructed and properly understood, i.e., as tentative candidates and not as dogmas, they can support and complement astrobiological explorations.

Though existent, the N = 1 problem is not as severe and insurmountable as Cleland seems to assume. A deep and sharp analysis of a reduced sample can be enough, at least in principle, to propose and scientifically validate the theoretical identity of a natural kind. The whole point lies in finding the right level of analysis to fix the identity criterion. The very favourite example of Cleland, i.e., water is H_2O, shows that once the right level of analysis is found, we do not need extraterrestrial samples of a natural kind to propose and validate a theoretical identity.

Today, thinking of possible samples of extraterrestrial water made of chemicals other than H_2O makes no sense to us, but this was not evident before setting the chemical criterion as the one fixing the natural kind of water. It was only when chemists established the identity between (earthly) water and H_2O that all other properties of familiar, earthly water (e.g., the typical three physical phases in which we find it) became interpretable as not essential, i.e., as something that, in principle, might vary in other corners of the universe. Before that, any of the different properties of water could be considered the universal criterion for its identification. For instance, in premodern conceptions, solvent power was considered an essential property of water, so different chemical compounds were grouped as simply sub-kinds of water: nitric acid as the so-called *aqua fortis* (strong water), and hydrochloric acid as *aqua regia* (royal water) (Roberts 1994; Cleland

and Chyba 2007). With that criterion, it would have made sense to wonder whether there were, across the universe, other kinds of water made of different, perhaps exotic chemical materials. Yet, science did not proceed that way because it chose a different level of analysis, the microstructural chemical one, to fix water's theoretical identity. Within that framework, the finding of the H_2O microstructure in earthly water was enough to fix the theoretical identity of the whole kind because, chosen as the identity mark, it immediately provided us with a universal identification criterion (being irrelevant whether the discovery of H_2O was made on Earth or somewhere else). Thus, today, we conduct astronomic explorations to see where there is water in space, not to see whether the universe hides forms of water made of something other than H_2O. If earthly water was good enough to fix a natural kind, why not earthly life?

We might think that the case of life is entirely different, but it is not. The N = 1 problem arises because earthly life (presumably) comes from only one common ancestor (LUCA), and it is reasonable to assume that there might be other independent origins and forms of life in the universe. But what if the case was different? Suppose we had living beings on Earth with distinct, independent origins (e.g., endowed with functional equivalents of enzymes or genetic material but made of entirely different compounds). Would that solve the problem of representative sampling? Against the set of all possible chemical variants in the universe, would an N = 2 or an N = 3 solve the problem? As the case of water teaches, the point is not whether we have diverse enough samples of a natural kind but rather whether we find the right level of analysis to fix the identity criterion for it. Without that, we lack a guide to tell which properties are to be considered as the identity mark of the kind and which are not. As I will show in the next sections, once we are equipped with the right level of analysis, which I argue in this case is the thermodynamic organisational one, earthly life proves to be enough to elaborate a theoretical identity for living beings. After all, if living beings constitute a natural kind, as I am assuming here they do, then the deep study of one sample, earthly or otherwise, at the right level of analysis should reveal the theoretical identity of the whole kind.

Now, suppose we concede to Cleland that the N = 1 problem is crucial and inescapable. What would that imply? Following Cleland's logic, it would mean that biology, until solving said problem, is not in a position to elaborate a universal theory of life. It would not mean, though, that there is no way to get a theoretical identity for living beings. The latter only follows if we assume, as Cleland seems to do, that biology and only biology can and should provide us with such a theoretical identity. This assumption, however, is neither forceful nor reflective of the practice of sciences in general. Scientific disciplines usually establish theoretical identities for the natural sub-kinds that are distinguishable within the general natural kind they study, not for this latter (i.e., not for the general kind they study). To illustrate, we expect chemistry to establish the theoretical identities of different chemical species and compounds, such as water (and others), not (necessarily) to establish what is a chemical entity in the first place as a general kind. It is rather subatomic, quantum physics that theoretically tells and explains what a chemical entity is in the first place.

Similarly, we can expect biology to tell us, for instance, why dolphins, despite their external appearance, are not fish. Biology can distinguish between mammals and non-mammals as sub-kinds of the general natural kind of living beings. But there is no forceful reason to demand that biology, and only biology, tells us what a living being is as a general kind. If living beings are a natural kind, then it makes sense to see them as a particular sub-kind of a more general kind and ask the science that deals with the latter what the distinctive properties of the biological sub-kind are.

If we follow this logic, the fact that current biology lacks the empirical and theoretical resources to develop a universal theory of life should not lead to the conclusion that we cannot (or should not) elaborate a theoretical identity for living beings. Cleland points out that the theoretical identity statement ‘water is H_2O’ represented an empirical discovery on the basis of a well-developed, mature and “widely accepted scientific theory for understanding and exploring a domain of natural phenomena” (2019b: 74). This is right. However, she makes a curious translation of this example to the case of life. In the case of water, the widely accepted scientific theory was the atomic-molecular theory, not an alleged mature, universal theory of water in particular. To establish the chemical nature of water, it was not necessary to have a universal theory of water or a well-established science of water (a waterology, so to speak). It was enough to have a well-developed general molecular theory. In the case of life, however, Cleland demands we should have a universal theory of life, an astronomically grounded biology, to elaborate, propose and validate a theoretical identity statement for living beings.

As I hope to show in the next section, living beings belong to the more general kind of dissipative structures. Therefore, it is not biology but thermodynamics -particularly the branch of far-from-equilibrium thermodynamics- that provides the general theory (criterion (iii) mentioned above) to understand the specificity of living beings as a natural kind. Thermodynamics is a mature scientific discipline with universal laws considered among the most robust and well-established in science. It should be, I submit, an excellent place to search for the theoretical identity of living beings.

Finally, from the more practical, methodological point of view, think of the following issue. Let us say we accept Cleland’s recommendation to wait until solving the N = 1 problem to propose a theoretical identity for living beings. When (or how) will we decide that our sampling of life in the universe is diverse and representative enough? How will we fix a precise point (in the seemingly endless exercise of sampling life in the universe) where the conditions to build a general theory of life are at last granted? What would be the “N,” considering the millions of galaxies and planets in the universe, that would satisfy our sampling? N = 100? N = 1,000? N = 1,000,000?

It seems more practical to join efforts and work, at the same time, on both the empirical and theoretical sides of the question. Astrobiological explorations and the elaboration of theoretical proposals, if well conducted, may feed and illuminate each other. A provisory but instrumentally good enough theoretical identity may provide a general guide to search for extraterrestrial samples of life. It may provide a normality criterion against which astrobiological explorations can interpret some data, as Cleland wants, as anomalous but genuinely indicative of life. Precisely because finding a satisfactory answer to the question “What are living beings’” is, as Cleland herself recognises, extremely complex, it sounds reasonable to work on both fronts of the problem rather than limiting the efforts to only one. In what follows, I will offer a candidate for a theoretical identity of living beings.

The proposal, of course, is offered not as a conclusive word but as a tentative answer to the problem of the theoretical identity of living beings, to wit, as a candidate to be evaluated, tested and reviewed in the light of theoretical and empirical advances. The spirit of the proposal is far from the spirit of venerable theoretical commitments such as the infamous “dogma” of molecular biology and other similar cases.

3. The thermodynamic autopoietic mark of living beings

In this section, I propose and explain the following theoretical identity statement: living beings are exo-autopoietic bodies. I argue that the statement, which is a thermodynamic reformulation of the autopoietic theory (Maturana 1975, 1981; Maturana and Varela 1980), adequately captures the ontological core that is common to all and only living beings.

Let me start by introducing the sense in which I will use the notions of autopoiesis and body in this proposal. Poiesis means production. Thus, poietic systems are systems that produce something or that perform some kind of production process. Production processes are those that build, compound or synthesise something, either by assembling parts or modifying the composition of a given structure. That which is produced in a poietic system, say cars in an automobile factory or chemical compounds in a chemical network, may or may not act as a constituent of the system itself. If it does, we speak of an autopoietic system (self-producing system). If it does not, we speak of an allopoietic system (systems that produce things other than themselves). The automobile factory, for instance, produces things (i.e., cars) that do not remain as constituents of the factory (they go out to the market), and it is, therefore, an allopoietic system. Autopoietic theory, in its classical formulation, asserts that living beings are molecular autopoietic systems because they produce, internally, many of the essential molecules and molecular structures that constitute them (Maturana 1975, 1981; Maturana and Varela 1980).

A body, on the other hand, is a discrete, physical object whose components keep spatial proximity in such a way that they remain as a unitary object. A simplified but useful way to think of a body is in terms of the physical phase. Gasses, or masses of gas, in general, do not constitute bodies as their component molecules diffuse in space (stars are a remarkable exception that we will comment on later). Solids, on the other hand, are typically bodies. This is just an approximation, as the key idea in the notion of a body, as I will use it here, is not the phase of the system (solid, liquid, gas, gel, etc.) but the maintenance of the spatial proximity of their components. Or, put another way, the resistance to the entropic disaggregation or spatial diffusion of them. We will discuss these notions in more detail when applied to the case of living beings. For now, let us step back to the general category of natural systems where living beings belong.

Natural systems can be categorised along a thermodynamic continuum that extends from an equilibrium condition (i.e., a condition of maximum entropy) to one far from it. Whether a system is at thermodynamic equilibrium, near to it, or far from it depends on several physicochemical factors, including the system's openness or closure to energy/matter exchange and its exposure to energetic fields or gradients. Using the equilibrium condition as a point of reference, we identify systems close to the equilibrium condition (near-equilibrium systems) and others far from it (far-from-equilibrium systems). Since the seminal works of Breuer (Elek and Müller 2013; Grandpierre, Chopra and Kafatos 2014) and Schrödinger (1944), it has been recognised that living beings belong to the group of far-from-equilibrium systems. In particular, these are known as dissipative structures (Nicoles and Prigogine 1977; Prigogine and Stengers 1984; Lineweaver 2006; England 2020).

Dissipative structures are open, far-from-equilibrium thermodynamic systems that self-organize as dynamic patterns under physicochemical gradient conditions. Typical examples of dissipative structures are Bènard cells, hurricanes, flames, and autocatalytic networks, among others. Different dissipative structures exhibit different kinds of constitutive dynamics. Some are constituted as chemical reactions (e.g., oscillatory chemical clocks, flames), some as purely thermal or mechanic patterns (e.g., convection patterns in Bènard cells, movement patterns of masses of air in hurricanes). They are called dissipative structures because they are ordered dynamic patterns (structures) formed and maintained through energy dissipative processes (hence dissipative). If we see energy dissipation without structure or structure without energy dissipation, then we are in the presence of something other than a dissipative structure. Even more, the mere concomitance of structure and dissipation is not enough to identify a system as a dissipative structure. A working car is a structure that processes and dissipates energy but is not a dissipative structure because its pieces and components are not formed and maintained through said dissipative process. Dissipative structures form and maintain *thanks* to the energetic and/or material flux they channel. The latter is, for them, not a facultative condition but a necessary one.

Now, energy dissipation is an entropic phenomenon, whereas forming and keeping an organised pattern or structure is a counter-entropic one. Insofar as dissipative structures are organised patterns of some kind, they all resist entropy in one dimension or another. Convection cells and hurricanes, for instance, are organised patterns of movement of molecules and masses of air, respectively, and therefore resist the entropic tendency to the free diffusion and homogeneous distribution of their components in space. On the other hand, autocatalytic networks are self-preserving patterns of chemical reactions and, therefore, resist the entropic tendency to chemical decay.

As Schrödinger (1944) and many after him have recognised, living beings are systems that locally resist entropy (England 2020). What has not been sufficiently recognised is the particular combination of entropic resistances that living beings perform. Here is where the autopoietic theory, particularly its embodied version, comes as a complement.

According to the autopoietic theory, living beings are systems that continuously produce their own components, i.e., autopoietic systems (Maturana 1975, 1981; Maturana and Varela 1980). Cells produce, through chemical reactions, the molecules, macromolecules and molecular structures that constitute them as such. Likewise, through chemical reactions and cell reproduction, organisms produce the components and structures (e.g., molecules, tissues, organs) that constitute them. As autopoietic systems, living beings resist the entropic tendency to the cessation of productive processes (chemical decay). But living beings are also bodies. They are organised as cohesive physical structures that behave as unitary objects (Villalobos and Razeto-Barry 2020). The key property here is the capacity to maintain the physical proximity of the components (against the entropic tendency to spatial diffusion or disorganisation) (Razeto-Barry 2012). The notion of an autopoietic body aims to capture this conjunction of properties (Villalobos 2020).

Living beings are systems that simultaneously resist, in a co-dependent way, both the cessation of their production processes and the tendency to the spatial diffusion or disorganisation of their components (Agmon 2020). Put in positive terms, living beings keep running their constitutive production processes while maintaining the spatial proximity of their components. In a cell, for instance, the proximity of the molecules allows them to enter into chemical reactions at the time that these chemical reactions produce and renovate the molecules that constitute and preserve said proximity. It is the circularity of these two thermodynamic conditions, i.e., resistance to spatial diffusion or disorganisation (body) and resistance to productive decay (autopoiesis), that characterises living beings as dissipative structures.

To illustrate the specificity of this condition, notice that living beings are not the only dissipative structures involving some sort of circular metaphysical dependency. Think of candle flames. A candle flame generates heat, which melts the wax, which exposes new portions of the wick, which feeds the flame that generates heat, and so on in a circular fashion. This circularity, however, does not involve the generation and preservation of a body. Candle flames are not bodies. In a flame, the chemical reactions' products (mainly carbon dioxide and vapour water) are constantly diffused at a high rate; they are not kept in spatial proximity. The more or less well-defined, luminous shape we usually identify as the flame is not a body; it is just the visible (light-emitting) part of the flux of exothermic, highly diffusive reactions that constitute the flame as a dissipative structure (Razeto-Barry and Ramos-Jiliberto 2013).

Notice also that living beings are not the only autopoietic systems. Autocatalytic networks are chemical systems that produce themselves, too. In an autocatalytic network, products of the chemical reactions act as catalysers of the chemical reactions that produce the components of the network. The difference, again, is that they are not bodies. In an autocatalytic network, the proximity of the components is kept by some container or recipient (e.g., a test tube), which is neither a part nor the product of the chemical network (Villalobos 2020).

The bodily nature of living beings needs to be correctly understood, that is, in terms of the integrated proximity of their components and not in terms of how (or by means of what kind of structure) such proximity is obtained. In unicellular living beings, the cell membrane is a key structure that keeps the proximity of the components because it acts, among other things, as a container, as a border against the diffusive forces of the molecules. However, this container structure is not a universal feature among living beings. In big multicellular living beings (e.g., mammals), the functional proximity of the components is kept not by means of a border or container but through the cohesive and adhesive forces of the histological/anatomical structures that constitute the organism (e.g., connective tissue, ligaments, fasciae, etc.) (Razeto-Barry and Ramos-Jiliberto 2013; Villalobos and Razeto-Barry 2020). In animals, the skin, which would be the equivalent of the cell membrane, is a complex organ with many vital functionalities. If you remove the totality of the skin of an animal, the animal will die after a while, but not because of a massive and rapid diffusion of the internal cells, tissues and organs. The skin has a key role in immunological protection, thermal regulation and other vital variables, but it is not a physical container to resist the diffusion of the internal components of the organism.

I have said that living beings are bodies that produce themselves, or, in other words, self-producing systems that are structured as bodies. But there is something more.

Living beings are bodies that realise their autopoiesis by feeding on the external environment, a thermodynamic feature that I propose to specify with the notion of an exo-autopoietic body. Living beings are open thermodynamic systems that take energy and matter from the environment and transform them, through productive processes, into their constitutive components. Why this specification? Think of stars. A star is a massive celestial body that keeps the proximity of its components thanks to the gravitational field generated by its own core. Core, where thermonuclear reactions forced by the gravitational field produce components that, with their mass, contribute to generating and preserving the very gravitational field that permits the proximity of the components that constitute the star. This circular organisation, where productive processes (i.e., thermonuclear reactions), enabled by the proximity of the elements, generate components that contribute in turn to maintaining the very proximity that enables the productive processes, corresponds essentially to the one of an autopoietic body. Where is the difference regarding living beings? The difference is that the star's autopoietic processes feed on internal resources rather than external ones. A star is a plasmatic body (a concentration of electrified gas) that consumes its own material resources to produce new components. It is, unlike the living being, an endo-autopoietic body.

Considering all these points, I propose to answer the question, "What are living beings?" with the formula, 'Living beings are exo-autopoietic bodies.'

In the next section, I will explain why and in which sense the formula has to be interpreted as a theoretical identity statement. But before that, a last point that needs to be mentioned regarding the notion of an exo-autopoietic body is its applicability criterion. An exo-autopoietic body is a dynamic entity, not a static one. It exists because of the productive processes it performs. In this sense, the notion of an exo-autopoietic body works as a theoretical criterion for identifying living beings in the present, i.e., while they are alive. A dead animal is, if you will, a biological entity insofar as we can study it from the point of view of biology in a broad sense. Still, it is not a living being since its autopoietic processes have ceased. Similarly, if you freeze a cell, you stop the exo-autopoietic processes (the cell's chemical reactions) and get a crystal body, which continues to be, if you will, a biological object for all practical purposes of biologists but not a living being. The frozen cell might return to life if carefully defrosted (Fuller, Lane and Benson 2019), but while frozen, it is a crystal, not an exo-autopoietic body, and, therefore, not a living being.

4. Living beings' theoretical identity

The formula 'living beings are exo-autopoietic bodies' is to be interpreted as a theoretical identity statement. This allows us to understand its theoretical status better, avoiding confusion and unnecessary problems. Theoretical identity statements are theoretical abstractions that aim to identify the ontological basis or core of a natural kind. This means two things. First, they do not aim to characterise the concrete manifestations of a natural kind exhaustively; that is, we do not ask them to be (necessarily) literal identities (Cleland 2019b). Second, they do not aim to provide a guide or criterion to establish neat and unambiguous spatiotemporal borders or cuts in continuous domains; we do not ask them to solve sorites like puzzles (Villalobos 2020). Let me illustrate, once again, with the example of the theoretical identity 'water is H_2O.'

First, following chemists, we accept that water is H_2O even when we know that real, concrete samples of water (e.g., tap water, a river) are never a chemically pure collection of H_2O molecules. They are always a mix of H_2O molecules and other compounds. Even chemically pure water samples are not a collection of only H_2O molecules since there is always a percentage of isotopic water molecules made of deuterium or tritium, plus ionic dissociative products such as $H+$, H_3O+ and $OH-$ (Hoefer and Martí 2019; Chang 2012). The statement 'water is H_2O' is, in this sense, an abstraction. It informs us about the ontological core that all and only water samples have in common, and we accept this, knowing that those samples are never uniquely H_2O. We do not ask scientific theoretical identities to be literal, concrete identities.

Similarly, when I hold that living beings are exo-autopoietic bodies, I do not mean that living beings are literally identical to, or exactly coincident with, the exo-autopoietic body they are. I propose a theoretical abstraction that informs us about the thermodynamic distinctiveness of living beings as natural systems, not an exhaustive characterisation of their realisation as concrete entities. In a concrete living being, only a portion of its constituent structures and processes correspond to an exo-autopoietic body. On the one hand, many of their vital processes are not poietic in nature (let alone autopoietic), and on the other, many of their structural components do not belong to their body in the specific thermodynamic sense I am using here. In a cell, for example, vital processes such as osmosis and ionic transportation are not poietic in nature (they move molecules in space but do not produce them). In animals, physiological processes such as gas transportation, membrane excitation and others are not autopoietic either. Even more, not all poietic processes that take place in living beings are autopoietic. In a neuron, for instance, the production and secretion of neurotransmitters is an allopoietic process. In animals, excretion is an allopoietic process too.

Regarding the body, in animals, for example, structures such as hair, hooves, and horns, among others, are parts of the organism's physical constitution as concrete entities. However, they are not parts of the animal's exo-autopoietic body as they play no role in preserving the proximity of the animal's cells, tissues, and organs. Not every component of the organism's physical (anatomical) body is a component of its exo-autopoietic body.

As with water and its theoretical identification with H_2O, concrete living beings are always more than exo-autopoietic bodies. However, their being exo-autopoietic bodies, I argue, is what identifies them as a natural kind.

Second, when chemistry proposes the theoretical identity statement ‘water is H_2O’, it is not meant to provide operational criteria to tell exactly when, at which precise point in time, something becomes or ceases to be water. Neither is it intended to offer empirical criteria to determine at which point in space a molecule (or collection of molecules) of water begins or ends. Asking for clear-cut borders in continuous domains generates what is known as sorites puzzles. Legitimate and entertaining as they may be, science does not work hard to propose theoretical identity statements for us to solve sorites puzzles. We might ask, exactly at which point in the interaction between two atoms of hydrogen and one of oxygen does water constitute? Precisely at which spatiotemporal point of their electrons’ interaction are we going to say that these atoms have formed the chemical bonds of a molecule of water? Legitimate as these questions may be, a guide to answer them is not what we expect from the theoretical statement ‘water is H_2O.’ Or, at any rate, the validity of the latter is not something we should measure based on how useful it is for solving them.

Similarly, the formula ‘living beings are exo-autopoietic bodies’ does not intend to provide us with operational criteria to tell when and where, precisely at which spatiotemporal point, a living being begins or ends as such. We might ask, how can we tell exactly where, in the anatomy of an organism, its exo-autopoietic body begins and ends? Or, when freezing a living cell, at what point can we say that the autopoietic processes have fully stopped and the cell has ceased to be alive? These questions' borderline cases and fuzzy boundaries, legitimate as they may be, are not problems we should expect the theoretical statement ‘living beings are exo-autopoietic bodies’ to solve. Or, at least, its validity should not be measured against that kind of problem.

5. The natural kind of living beings

In the previous section, I argued that the formula ‘living beings are exo-autopoietic bodies’ must be interpreted as a theoretical identity statement—that is, as a theoretical abstraction. In this section, I want to focus on the formula's specific content (not in its form). Why is a thermodynamic version of the autopoietic theory (or, conversely, an autopoietic version of the thermodynamic approach to life) a good candidate for a theoretical identity of living beings?

Theoretical identities require a certain degree of abstraction since they target a natural kind in its generality, not particular cases. However, said abstraction cannot be excessive. Otherwise, we miss the required specificity to identify the natural kind as distinct from others. For instance, saying that water is a combination of hydrogen and oxygen atoms without specifying the exact proportion of the combination is too general (hydrogen peroxide and other hydrogen polyoxides that are not water would qualify as water). Saying that water is a combination of two atoms of the same kind plus one of another, without specifying the elements, is again too abstract (many compounds with the ratio 2:1 that are not water would qualify as water). In the opposite direction, saying that water is only orthowater (i.e., a subspecies of H_2O molecules where the spins of the protons are aligned) is too narrow as parawater (H_2O molecules where the spins of the protons are not aligned) counts equally well as the microstructure that explains the characteristic properties of water (Hoefer and Martí 2019; Häggqvist and Wikforss 2018).

This delicate balance between generality and specificity is critical in the case of living beings. Traditional approaches to life fail to find this balance and, consequently, fail to pass either the exclusivity test (i.e., to indicate a property or conjunction of properties only living beings have) or the universality test (i.e., to indicate a property or conjunction of properties all living beings have). This failure is not because of their form—that is, because they come as definitions—but because of their content.

I argue that the idea of an exo-autopoietic body provides the right level of abstraction (the right balance between generality and specificity) required to identify living beings as a natural kind. Let me illustrate through some examples and comparisons.

The traditional biochemical approach identifies a set of chemical compounds as distinctive life markers, typically a carbon-based substrate with specific molecules such as nucleic acids and enzymes. This sounds reasonable as long as we remain focused on earthly life. However, the commitment to a particular set of chemical compounds risks failing the universality test when we consider the possible chemical variants of hitherto unknown life forms across the universe (see, for instance, Schulze-Makuch and Irwin 2008 for the plausibility of non-carbon-based forms of life). The notion of an exo-autopoietic body, instead, provides a better level of abstraction because it refers to a determined thermodynamic organisation but does not commit to any particular chemical (carbon-based or other) realisation.

The physiological approach identifies living beings with a series of characteristic processes, such as eating, digesting, excreting, moving, etc. This approach fails to meet the exclusivity test as, under this criterion, some trivial human-made machines could qualify as living beings. An automobile, for instance, can be said to eat, digest, excrete, move, and react to external stimuli (Sagan 1970). The idea of an exo-autopoietic body blocks this counterexample as it requires the system to produce its own components, where these components constitute the very body of the system. A functioning automobile performs combustion chemical reactions on the inside, and it is a body. Still, from those chemical processes, no piece of the car (as a mechanical structure, a concrete body) is produced.

Growth is another ability usually considered exclusive to and universal in living beings. However, it fails both the exclusivity and the universality tests. First, inorganic crystals, under the right conditions and provided with enough ion sources, can grow, too (Schulze-Makuch and Irwin 2008). Second, growing is not an ever-present process in living beings. Growing may be observable in some specific phases of the life cycle. However, in most observation windows, growing is absent. Should we judge an adult animal not being alive just because it is not growing? The notion of an exo-autopoietic body makes no commitment to growing and thus remains valid, disregarding whether growing is present or not.

The evolutionary approach points out the ability to reproduce and evolve by natural selection as the distinctive mark of living beings. This criterion, however, fails the universality test as there are living beings that do not reproduce. For instance, worker bees and some hybrid animals, such as mules, do not reproduce. Reproduction and the ability to evolve are not, therefore, universal properties in living beings. The notion of an exo-autopoietic body does not presuppose reproduction and evolution; therefore, it remains valid for living beings that do not reproduce and evolve.

The classical thermodynamic approach to life, after Schrödinger (1944), highlights living beings' ability to generate internal order (to decrease entropy) at the expense of increasing external disorder or entropy. This approach, however, proves to be too general and vulnerable to counterexamples such as mineral crystals (e.g., quartz or diamond), which create order by extracting energy/matter from the environment (Schulze-Makuch and Irwin 2008; Sagan 1970). As I exposed before, living beings' peculiarity is not that they locally decrease entropy (all dissipative structures do that) but the particular form they do it: resisting spatial diffusion or disorganisation and production decay in a co-dependent way while energetically feeding on the external environment. That is what an exo-autopoietic body is. Thermodynamic approaches that do not take this autopoietic specification fail because they permit every dissipative structure (e.g., hurricanes, convection cells) to be qualified as a living being (see, for example, Lineweaver 2006).

Traditional systemic approaches to life, such as the classical autopoietic theory, are too abstract and fail to pass the exclusivity test. In its original formulation, the autopoietic theory talked of living beings in terms of production processes, paying little attention to their physical conditions of realisation (see Maturana and Varela 1980). This led to interpretations in which formal, computational models were thought to meet the autopoietic criteria to be considered alive (Cleland 2019b). Also, macro or supra-biological systems such as societies and ecological trophic networks were considered higher-order autopoietic systems and, therefore, forms of life (Lovelock 2000; Margulis and Sagan 1995). Later, the theory made explicit the specification that living beings are molecular autopoietic systems (Maturana 2011, 2019). Yet, as we saw before, even under this specification, we find systems such as autocatalytic networks that qualify as living beings (Villalobos 2020). The notion of an exo-autopoietic body shows that living beings' peculiarity is not their autopoietic organisation but the particular thermodynamic incarnation it takes. It reduces to an adequate level of concreteness the high degree of abstraction of classical autopoietic theory.

6. Conclusion

Theoretical identity statements seem to represent, better than definitions, the scientific practice of identifying a natural kind by means of theoretical abstractions that reveal its bases or deep structure. It is this methodological gesture that, I argue, is worth replicating with the question of living beings.

Some, however, think that the gesture cannot (and should not be) replicated because biology lacks a universal (or general enough) theory of life due to the limited (only earthly) sample it works on. In particular, Cleland and Chyba recommend focusing on astrobiological explorations to fill this biology's empirical gap instead of proposing premature theoretical identities for living beings.

I propose an alternative interpretation in which we ask not biology but a more basic scientific branch, far-from-equilibrium thermodynamics, for the identity of living beings as a natural kind. Complementing the thermodynamic approach with the insights of autopoietic theory, I claim that the question "What are living beings" can be satisfactorily answered with the theoretical identity statement 'Living beings are exo-autopoietic bodies.'

I offer this theoretical candidate, tentative and open to scientific disconfirmation, not to replace but to complement the astrobiological explorations recommended by Cleland and Chyba.

References

Agmon E (2020) Deriving the bodily grounding of living beings with molecular autopoiesis. Adaptive Behavior 28 (1): 35–36. https://doi.org/10.1177/1059712319830645

Bedau M, Cleland, CE (eds) (2010) The nature of life: Classical and contemporary perspectives from philosophy and science. Cambridge University Press. https://doi.org/10.1017/CBO9780511730191

Bich L, Green S (2018) Is defining life pointless? Operational definitions at the frontiers of biology. Synthese 195: 3919–3946. https://doi.org/10.1007/s11229-017-1397-9

Chang H (2012) Is water H_2O? Evidence, realism and pluralism. Springer, Netherlands

Cleland CE (2019a) Moving beyond definitions in the search for extraterrestrial life. Astrobiology 19 (6): 722-729. https://doi.org/10.1089/ast.2018.1980

Cleland CE (2019b) The quest for a universal theory of life: Searching for life as we don't know it. Cambridge University Press. https://doi.org/10.1017/9781139046893

Cleland C E (2012) Life without definitions. Synthese 185:125–144. https://doi.org/10.1007/s11229-011-9879-7

Cleland CE (2006) Understanding the nature of life: A matter of definition or theory? In: Seckbach J (ed) Life as we know it. Springer, Dordrecht, pp 589–600

Cleland CE, Chyba CF (2002) Defining 'life.' Orig. Life Evol Biosph 32: 397–393. https://doi.org/10.1023/A:1020503324273

Cleland CE, Chyba CF (2007) Does ‘life’ have a definition? In: Sullivan III WT, Baross JA (eds) Planets and Life: The Emerging Science of Astrobiology. Cambridge University Press, Cambridge, pp 119–131. https://doi.org/10.1017/CBO9780511730191.032

Elek G, Müller M (2013) The living matter according to Ervin Bauer (1890–1938) (on the 75th anniversary of his tragic death) (History). Acta Physiologica Hungarica 100(1): 124–132. https://doi.org/10.1556/APhysiol.99.2012.006

England J (2020) Every Life Is on Fire: How Thermodynamics Explains the Origin of Living Things. Basic Books, NY, USA

Fuller BJ, Lane N, Benson EE (eds) (2019) Life in the frozen state. CRC Press, Boca Raton, FL

Grandpierre A, Chopra D, Kafatos MC (2014) The universal principle of biology: Determinism, quantum physics and spontaneity. NeuroQuantology 12 (3): 364-373. DOI: 10.14704/nq.2014.12.3.747

Häggqvist S, Wikforss Å (2018) Natural kinds and natural kind terms: Myth and reality. The British Journal for the Philosophy of Science 69 (4): 911-933. https://doi.org/10.1093/bjps/axw041

Hermida M (2016) Life on Earth is an individual. Theory Biosci. 135: 37–44. https://doi.org/10.1007/s12064-016-0221-2

Hoefer C, Martí G (2019) Water has a microstructural essence after all. Euro Jnl Phil Sci 9, 12. https://doi.org/10.1007/s13194-018-0236-2

Kolb V (2018) Handbook of Astrobiology. CRC Press, Boca Raton, FL

Lineweaver CH (2006) We have not detected extraterrestrial life, or have we? In: Seckbach J (ed) Life as we know it. Springer, Dordrecht, pp 445-457

Lovelock J E (2000) The ages of Gaia: A biography of our living Earth. OUP, Oxford

Machery E (2012) Why I stopped worrying about the definition of life…and why should as well. Synthese, 185, 145-164. https://doi.org/10.1007/s11229-011-9880-1

Margulis L, Sagan D (1995) What is life? University of California Press, Berkeley, CA

Mariscal C, Doolittle WF (2020) Life and life only: a radical alternative to life definitionism. Synthese 197: 2975–2989. https://doi.org/10.1007/s11229-018-1852-2

Maturana H (2020) Reflections in relation to the article of Villalobos and Razeto. Adaptive Behavior 28(1): 15-17. https://doi.org/10.1177/1059712319841740

Maturana H (2011) Ultrastability ... autopoiesis? Reflective response to Tom Froese and John Stewart. Cybernetics & Human Knowing 18 (1-2): 143–152.

Maturana H (1975) The organization of the living: A theory of the living organization. Int J Man-Mach Stud 7: 313–332.

Maturana H (1981) Autopoiesis. In: Zeleny M (ed) Autopoiesis: A Theory of Living Organization. North Holland, NY, USA, pp 21–33

Maturana H, Varela F (1980) Autopoiesis and Cognition: The Realization of the Living. Kluwer Academic Publishers, Dordrecht, The Netherlands

Mix LJ (2015) Defending definitions of life. Astrobiology 15(1). https://doi.org/10.1089/ast.2014.1191

Nicolis G, Prigogine I (1977) Self-organization in nonequilibrium systems: From dissipative structure to order through fluctuations. Wiley, New York

Pályi G, Zucchi C, Caglioti L (2002) Introduction: definitions of life. In: Pályi G, Zucchi C, Caglioti L (eds) Fundamentals of Life. Elsevier, New York, pp 2–13

Prigogine I, Stengers I (1984) Order out of Chaos: Man's New Dialogue with Nature. Bantam Books, NY, USA

Razeto-Barry P (2012) Autopoiesis 40 years later: A review and a reformulation. Origins of Life and Evolution of Biospheres 42: 543–567. https://doi.org/10.1007/s11084-012-9297-y

Razeto-Barry P, Ramos-Jiliberto R (2013) ¿Qué es autopoiesis? [What is autopoiesis?]. In: Razeto-Barry P, Ramos-Jiliberto R (eds) Autopoiesis: Un Concepto Vivo [Autopoiesis: A Living Concept]. Editorial Nueva Civilización, Santiago, Chile, pp 27–57

Roberts G (1994) The Mirror of Alchemy: Alchemical Ideas in Images, Manuscripts and Books. Toronto: University of Toronto Press.

Sagan C (1970) Life. In: Encyclopedia Britannica. Encyclopedia Britannica Incorporated, Chicago, pp 1083-1083A.

Schulze-Makuch D, Irwin LN (2008) Life in the universe. Expectations and Constraints. Springer-Verlag, Berlin Heidelberg

Schrödinger E (1944) What is life? Cambridge University Press, Cambridge

Trifonov E N (2011) Vocabulary of definitions of life suggests a definition. J Biomol Struct Dynamics 29(2): 259–266. https://doi.org/10.1080/073911011010524992

Tsokolov SA (2009) Why is the definition of life so elusive? Epistemological considerations. Astrobiology 9(4). https://doi.org/10.1089/ast.2007.0201

Villalobos M. (2020) Living beings as autopoietic bodies. Adaptive Behavior 28(1): 51–58. https://doi.org/10.1177/1059712319879747

Villalobos M, Razeto-Barry P (2020) Are living beings extended autopoietic systems? An embodied reply. Adaptive Behavior 28(1): 3–13. https://doi.org/10.1177/1059712318823723