Naturalized Teleology: Cybernetics, Organization, and Purpose

Carl B. Sachs

csachs@marymount.edu

School of Humanities in Marymount University

Abstract: The rise of mechanistic science in the 17th century helped give rise to a heated debate about whether teleology – the appearance of purposive activity in life and in mind – could be naturalized. At issue here were both what is meant by “teleology” as well as what is meant “nature”. I shall examine a specific episode in the history of this debate in the 20th century with the rise of cybernetics: the science of seemingly “self-controlled” systems. Against cybernetics, Hans Jonas argued that cybernetics failed as a naturalistic theory of teleology and that the reality of teleology is grounded in phenomenology, not in scientific explanations. I shall argue that Jonas was correct to criticize cybernetics but that contemporary work in biological organization succeeds where cybernetics failed. I will then turn to contemporary uses of Jonas’s phenomenology in enactivism and argue that Jonas’s phenomenology should be avoided by enactivism as a scientific research program, but that it remains open whether enactivism as a philosophy of nature should also avoid Jonas.

Keywords: teleology – autopoiesis – Hans Jonas – cybernetics -- naturalism

0. Introduction

It has become a familiar theme that the rise of mathematical mechanistic physics in the 17th century and its explanatory success across a wide range of phenomena put immense pressure on the question as to whether the concepts of purpose, intention, or even value had any legitimate role within a scientific worldview. Whether such concepts, seemingly indispensable for our ordinary self-conception, could be understood in naturalistic terms, or if naturalism required relegating such notions to the ontological dust-bin along with angels and ghosts, preoccupied many of the greatest philosophers and scientists since the Scientific Revolution began. The dialectic of mechanism vs teleology, including whether naturalism must be mechanistic and therefore cannot accommodate teleology and whether teleology could be or should be naturalized, is one of the most enduring themes of the Enlightenment and its aftermath, up to and including the present-day.

In what follows, I propose to explore this wide-sweeping problematic through a close examination of a specific moment of this dialectic: the critique of cybernetics by the bio-phenomenologist Hans Jonas. I have chosen this emphasis for three reasons. Firstly, the early cyberneticists explicitly understood their project as the construction of “teleological mechanisms”. Not only were they well-informed about the dialectic of mechanism and teleology, but they had considered reasons for why they thought that they had adequately resolved it (even if we today might think that they were mistaken).[[1]](#footnote-1) Secondly, the legacy of cybernetics has continued to shape the subsequent discourse about what it would mean to naturalize teleology. If this is less evident, I suggest that it is because we have largely forgotten what cybernetics was. Yet early cybernetics was in many respects an ancestor of recent prominent attempts to naturalize teleology, including autopoiesis, biological autonomy, or Deacon’s biosemiotics – though this is not to ignore or minimize the very real differences between those approaches. No history of autopoiesis or Deacon’s biosemiotics should ignore the importance of second-order cyberneticists Humberto Maturana and Heinz von Foerster, or the fact that in Terry Deacon’s *Incomplete Nature*, one of the most often-quoted philosophers is the second-order cyberneticist Gregory Bateson. Thirdly, Hans Jonas’s critique of cybernetics is of interest because Jonas is one of the most commonly cited phenomenologists in this branch of inquiry, especially in light of how his bio-phenomenology influenced the development of enactivism beginning with Weber and Varela (2002). Thus, contemporary developments in autopoiesis, biological autonomy, and enactivism are partly descended from cybernetics and from bio-phenomenology.

To begin this examination, I first reconstruct what cybernetics was at the time of Jonas’s critique and what his critique involved (§1). Jonas makes both a specific claim about cybernetics and far more generic claim about teleological realism, that is, the position that teleology exists, rather than being a mere fiction or heuristic. With regard to cybernetics, Jonas claims that it attempted to naturalize teleology in terms of circular causality, but this fails because it cannot account for the difference in kind between the purposiveness of artifacts and the purposiveness of organisms. In a separate line of argument, Jonas suggests that there is no room within a scientific worldview for teleology, but that teleological realism can be grounded in phenomenological considerations. As I read Jonas, we should look to phenomenology for the justification of teleological realism precisely because teleology cannot be accommodated by the natural sciences and hence cannot be naturalized.

I will then turn to a recent attempt to naturalize teleology that stand a chance of succeeding where cybernetics fails: the organizational approach (henceforth the OA) developed principally by Alvaro Moreno, Maël Montévil, Matteo Mossio, and Leonardo Bich (§2). I shall argue that the OA significantly advances beyond cybernetics in the following sense: *pace* Jonas, we can explain the “needful freedom” of the organism in terms of the circular causality of its organization. Whereas Jonas is right that the kind of circular causality theorized by cybernetics cannot account for the reality of biological teleology, the OA theorizes an importantly different kind of circular causality.

However, Jonas’s second point – that the reality of teleology must be based on phenomenology -- has important implications for contemporary philosophy of mind that are less easily resolved. It is this criticism that led Weber and Varela (2002) to urge that autopoietic enactivism converges with Jonasian biophenomenology. The central issue here is whether the status of teleology as a scientific concept can afford to lean upon phenomenological considerations. Here I shall argue that while a comprehensive philosophy of mind should not ignore phenomenology, enactivism as a scientific research program should follow the OA as a conceptualization of teleology that does not depend upon phenomenology. Hence Jonas’s phenomenological argument for the reality of teleology should be used very carefully by enactivists, if not avoided entirely (§3). I conclude by considering whether or not we should say that biological organization has successfully naturalized teleology depends just as much on what we mean by “naturalization” as it does on what we mean by “teleology” (§4)

1. Hans Jonas’s Critique of Cybernetics

Jonas first articulated his criticism of cybernetics in 1952, four years after the publication of Wiener’s landmark text introduced the word “cybernetics” to the reading public.[[2]](#footnote-2) Though Jonas cautiously acknowledges that a philosopher should leave to specialists questions about the legitimacy of cybernetics as a special science, it nevertheless becomes a philosophical question when cybernetics is looked to as solving philosophical problems or as playing a super-scientific role in the organization of knowledge:

Now is offered, for the first time, a mechanistic model which, it is claimed, applies to material and mental phenomena at once, not equivalently but identically – that is, without involving passage from field to field. This would indeed mean an overcoming of the dualism which classical materialism had left in possession by default: for the first time since Aristotelianism we would have a unified doctrine, or at least a unified conceptual scheme, for the representation of reality. Needless to say, this would be of the utmost philosophical importance – and the spokesmen of cybernetics are not restrained by any timidity from pointing out these implications in very explicit statements. (Jonas 2001: 110-111)

As Jonas sees it, there is an important difference between what could be called *scientific cybernetics* and *philosophical cybernetics*. Whereas scientific cybernetics is limited to the principles of information, signal, noise, and feedback in electronic engineering, philosophical cybernetics attempts to demonstrate how such concepts could be used, not analogically or metaphorically but univocally, to explain animal behavior and even abstract thought. If philosophical cybernetics were successful, Jonas realizes, we would have begun to heal the rift between the natural sciences and human sciences that began with Descartes’s insistence that a mechanistic mathematical explanation of the world stopped at the human mind. The popularizing statements of 1940s cyberneticists such as Norbert Wiener, Lawrence Frank, and F.C.S. Northrop show that Jonas was not exaggerating about the grandiosity of the early cybernetic vision; as Galison (1998) points out, cybernetics was seen by its foremost theorists and advocates as contributing to and building upon the Unity of Science movement.

To better understand Jonas’s critique of philosophical cybernetics, it should be emphasized that the early cyberneticists explicitly claimed that their interest was in “teleological mechanisms.” Prior to Heinz von Foerster’s suggestion that “cybernetics” be used to describe the Macy Conferences, the Macy Conferences were called the Macy Conferences in Circular Causality and Teleological Mechanisms (Dupuy 2009). Furthermore, a 1948 issue of *Annals of the New York Academy of Sciences* was entitled “teleological mechanisms” and contained papers by Wiener and McCulloch. The guiding idea was that examination of circular causality could explain the phenomena that philosophers had classified as teleological – above all, life and mind. As Lawrence Frank put it in his Foreword to the *Annals:*

The concept of teleological mechanisms, however it may be expressedin different terms, may be viewed as an attempt to escape from theseolder mechanistic formulations that now appear inadequate, and toprovide new and more fruitful conceptions and more effective methodologiesfor studying self-regulating processes, self-orientating systems andorganisms, and self-directing personalities. But these new concepts carryno psychic or vitalistic assumptions, nor do they imply that any mysterious,supernatural powers or psychic forces or final causes are operatingthe system or guiding the organism-personality. The idea of purposivebehavior is not a regressive movement to an earlier stage in the history ofideas, but a forward movement toward a more effective conception of the problems we face today. (Frank 1948: 191)

As Frank sees it, the goal of modern science is to replace the old conception of mechanism that privileged explanation in terms of analyses into constituent parts and statistical averages over large samples with a new conception of organisms as wholes that are different from the sums of their parts, the importance of configuration or organization in the study of self-directed, purposive systems. This conception, moreover, is held to be of supreme usefulness for addressing the social, economic, and political problems of a world recovering from the Second World War. Seen in this light, there is good reason why Hans Jonas’s “search for the good after Auschwitz” would have led him to undertake a sustained critical examination of philosophical cybernetics.

The crux of Jonas’s critique of cybernetics is that it fails as a theory of biological purpose because it has no concept of need, concern, or interest. Jonas thinks that life *as such* is distinguished by what he calls the “needful freedom” (Jonas 2001: 80) of the organism: to be alive is to be simultaneously independent of the environment and dependent upon it. But what justifies or grounds this ontological commitment? Here Jonas’s methodology is phenomenological or bio-phenomenological: “life can only be known by life” (ibid., p. 91). It is only because *we*, as phenomenological observers, are ourselves alive that we can correctly label, classify, and describe anything else as also being alive. That is, it is only because we have a primordial experience of ourselves as beings of needful freedom that we can correctly label, classify, and describe anything else – a lion, a starfish, a maple tree – as a fellow being of needful freedom.[[3]](#footnote-3)

The crux of Jonas’s criticism is that teleological mechanisms can never give us genuine insight into the nature of living organisms, into the ontology of life. Yet that is precisely what the early cyberneticists thought they were doing (among other things): explaining the nature of life in scientific terms by explaining the communication principles that life shares with feedback-governed machines. Thus Wiener, in his 1948 *Cybernetics*, insisted that cybernetics studies ‘communication engineering’ as such, ‘whether in the metal or in the flesh’, since it is best understood as

the newer study of automata, whether in the metal or in the flesh, is a branch of communication engineering, and its cardinal notions are those of message, amount of disturbance or ‘noise’ – a term taken over from the telephone engineer – quantity of information, coding technique, and the like. In such a theory, we deal with automata effectively coupled to the external world, not merely by their energy flow, their metabolism, but also by a flow of impressions, of incoming messages, and of the actions of outgoing messages. (Wiener 1948: 42)

Of particular relevance to Jonas’s critique is how Wiener conceptualized the activity of the central nervous system in terms of circular causation:

 The central nervous system no longer appears as a self-contained organ, receiving inputs from the senses and discharging into the muscles. On the contrary, some of its most characteristic activities are explicable only as circular processes, emerging from the nervous system into the muscles, and re-entering the nervous system through the sense organs, whether they be proprioceptors or organs of the special senses. (Wiener 1948: 8)

and

[F]or effective action on the outer world it is not only essential that we possess good effectors, but that the performance of these effectors be properly monitored back to the central nervous system, and that the readings of these monitors be properly combined with the other information coming in from the sense organs to produce a properly proportioned output to the effectors. (Wiener 1948: 96)

Even if, however, the basic contention of cybernetics was granted – the dynamic causal loops between organism and environment have the same formal structure as those of guided missiles – something of fundamental importance would be omitted. As Jonas sees it, circular causation and feedback loops are at best a necessary and never a sufficient condition of biological teleology, because they cannot explain that animals act purposively because of what they need or what is of concern or interest to them. Need, concern, and interest are concepts, Jonas contends, that cannot be explained in mechanistic terms, no matter how “teleological” the mechanisms are.

To make this point clear, Jonas says, suppose we begin with a thought-experiment of a human pilot guiding a torpedo. In this case the purpose belongs to the pilot, not to the instrument that she uses to execute her purpose. If we replace the pilot with a regulator that coordinates transducers and effectors, that regulator does not have intrinsic purposiveness. As a mechanical part it is as unable to resist increased entropy as is the rest of the system (pp. 117-118). Nor would it make sense, Jonas thinks, for purposiveness to be suddenly transferred to the torpedo as a whole. The torpedo cannot be anything other than an instrument for realizing someone’s purpose, which they have in light of how they, as a concrete individual, conceive of their ultimate concern, their conception of what is valuable and worthwhile in human life. Thus, cybernetics fails as a theory of biological purposive behavior; need, interest, concern, or value cannot be explained in terms of circular causation and feedback loops. Something more is needed, and no concept of teleological mechanisms can provide it. The “needful freedom” of the organism – its precarious status of dependence on and independence from the environment – cannot be understood in mechanistic terms, and more precisely in terms of increased entropy production.

If the needful freedom of the organism cannot be understood in mechanistic terms, then how can be it understood? In his “Is God a Mathematician?” Jonas invokes the idea of an omniscient observer knowledgeable of all (and only) the laws of physics. Could such a being discern the difference between life and nonlife. Jonas thinks the answer is ‘no’:

Every happening involved, and consequently their sum total in each instant, can and thus must be accurately accounted for on the lines of the general schemes underlying the mathematical-mechanical world-picture … the apparent sameness and individuality of the organic whole will resolve itself … into the network of all physical lines which converge here and now from the universe … and all the features of a self-related autonomous entity would, in the end, appear as purely phenomenal, that is, fictious. (Jonas 2001: 78)

Since an omniscient mathematical physicist would be limited to a mechanistic world-view, the reality of individual organisms would not show up for it – it would see atoms and molecules as only temporarily organized into organisms, and would be unable to see a difference that makes a difference between a unified biological individual and a mere temporary aggregate. The reason why we know that there is such a difference is because we experience that difference in ourselves:

here, being living bodies ourselves, we happen to have inside knowledge. On the strength of the immediate testimony of our bodies *we* are able to say what no disembodied onlooker would have cause for saying … in living things, nature springs an ontological surprise in which the world-accident of terrestrial conditions brings to light an entirely new possibility of being: systems of matter that are unities of a manifold … in virtue of themselves, for the sake of themselves, and continually sustained by themselves. (ibid.: 79).

We experience ourselves as living, autonomous beings in a way that the omniscient mathematical physicist could never predict or conceptualize; relative to a mechanistic world-view, life is an ‘ontological surprise’. It is because the scientific world-view is committed to mechanistic explanations that teleological realism must be grounded in a different form of understanding that that of modern science: it is grounded in phenomenology.

We can therefore summarize Jonas’s critique of cybernetics as having two prongs: that cybernetics fails as a mechanistic theory of biological teleology because the needful freedom of the organism cannot be explained in terms of circular causality and that teleological realism is grounded in phenomenology, not in science. I will now turn to recent work in theoretical biology to show that in one important sense, Jonas was wrong: we can understand the needful freedom of the organism in terms of circular causality just as long as we are careful to distinguish organizational closure from thermodynamic openness (§2). I will then use the OA to question the need for Jonasian phenomenology in enactivism as scientific research program and as philosophy of nature (§3).

2. The Organizational Approach to Naturalizing Teleology

The organizational approach (OA) to naturalizing teleology is a relatively new development in theoretical biology that intends to naturalize teleology in terms of a specific kind of formal structure distinctive of living things.[[4]](#footnote-4) In brief, it holds that life is distinguished from non-life by two distinct and complementary structural features: organizational closure and thermodynamic openness. According to the OA, thermodynamic openness and organizational closure are individually necessary and jointly sufficient for the needful freedom of life which explains teleology. Hence, we can understand teleology to be an intrinsic actual property of living things; we can accept teleology as a constitutive concept in biology without any anxiety about opening the door to vitalism or the supernatural. In that regard the OA purports to succeed where cybernetics failed (*pace* Lawrence Frank). In this section, I will first briefly describe the central concepts of the OA and show how the OA explains the needful freedom of living things in terms of circular causation. Doing so will allow me to explain why the OA succeeds where cybernetics fails.

The OA begins with the conception of biological systems as distinct from other kinds of physical systems by virtue of being *self-determining systems*. A system is self-determining just in case the activity of the system contributes to the generation and maintenance of the capacity of the system to resist entropy. This captures one important aspect of what makes life an ontological surprise relative to the laws of thermodynamics: biological systems are localized regions of entropy decrease. But it is also crucial that entropy decrease is not just something that happens to biological systems, as a result of favorable circumstances: the localized decrease in entropy is caused by what the system itself is doing with regard to the larger system in which it is embedded.

The OA approach then defines self-determination in terms of *the closure of constraints*. Systems are self-determining if “they generate and maintain a set of structures acting as constraints which, by harnessing and channelling the processes and reactions occurring in the system, contribute to sustain each other, and then the system itself” (Moreno and Mossio 2015: xxix). For this reason, we next need to provide a rigorous account of what they call, following Varela, “closure”. In abstract terms, closure is what allows biological systems to be self-determining: a system exhibits organizational closure if what the system does is determined, at least to some extent, by organizational dynamics of the system itself. At the same time, however, the persistence of organizational closure requires that the system is constantly interacting with its environment for both extracting energy and material as well as eliminating heat and waste. Moreno and Mossio aim to show that a fully adequate account of self-determination requires “two distinct, although closely interdependent, regimes of causation: an *open* regime of thermodynamic processes and reactions, and a *closed* regime of dependence between components working as constraints” (ibid.: 3; emphasis original).

*Pace* Varela, Moreno and Mossio understand closure not as a closure of processes but rather as “a specific kind of mutual dependence between a set of entities having the status of *constraints* within a system” (ibid.: 5; emphasis original). Constraints, in turn, are understood in terms of causes that reduce the degrees of freedom of the system in which they act. The key difference between physico-chemical systems and biological systems is that “thermodynamic flow is channelled and harnessed by a set of constraints in such a way as to realize mutual dependence between these constraints” (*ibid.*). In other words, what distinguishes living things from nonliving things is that living things are organized in such a way that they exploit energy gradients so as to realize an interlocking system of constraints, each of them reduces the degrees of freedom of the system. As a consequence, the system *as a whole* achieves a kind of self-determination precisely because the system tends to behave in ways that will generate and maintain these constraints. Importantly, the closure is at the level of constraints, and not at the level of thermodynamic processes; Moreno and Mossio emphasize that biological systems are both open and closed, in two distinct causal regimes – thermodynamic flows on the one hand (openness), causal constraints on the other (closure). As they see it, it is the interdependence between these conceptually distinct kinds of causal regimes that has been missing from previous attempts to specify the organizational approach.[[5]](#footnote-5)

To clarify how the organizational approach is supposed to work, it is helpful to consider how the OA would draw the distinction between biological autonomous systems and dissipative or self-organizing systems. Both involve a global structure is maintained at far-from-thermodynamic equilibrium with its environment. Classical examples of spontaneously self-organizing systems would include Bénand cells, flames, hurricanes, and Belousov-Zhabotinsky reactions. These systems are self-organizing in the following sense: the system involves a system-level constraint that determines the behavior of the constituents of the system so that the constituents act in ways that maintain the existence of the system as a whole. As a simple example, a burning candle produces heat that melts the wax that can then travel up the wick for combustion.

This is distinct from self-determining systems, however, because self-organizing systems are limited by boundary conditions that they themselves do not affect. Closure, by contrast, “involves the *takeover of (some of) the boundary conditions* required for the maintenance of the system” (17; emphasis original). Put otherwise, Bénand cells are self-organizing because the hexagonal structures tend to constrain the behavior of water molecules, but not self-determining because the Bénand cells play no role in determining the boundary conditions of the system, such as the source of heat. Biological cells are self-determining, and not just self-organizing, because the activity of the system contributes to the boundary conditions of the system, e.g., ensuring the continued function of the cell membrane, locating and absorbing sources of chemical energy, eliminating otherwise toxic waste products, etc.[[6]](#footnote-6)

It is now crucial to specify in more formal terms what Moreno and Mossio mean by closure. In their terms, closure requires “an organisation in which each constraint is involved in at least two different dependence relationships in which it plays the role of enabling and dependent constraint, respectively” (21); each constraint is both *constraining* and *constrained by*. If all constraints are both constraining and constrained by, the totality of all constraints is self-constraining. If the totality of all constraints is self-constraining, then the system realizes closure. In other words, one of the necessary conditions of a self-determining system is that, as closed system, it is organized with *mutually constraining* constraints.[[7]](#footnote-7)

Closure is, however, only one of the necessary conditions for biological autonomy. The other necessary condition is thermodynamic openness: the system must be exposed to energy gradients that can be harnessed by the system of constraints to perform work that generates and maintains that system. This is how, according to the OA, life manages to bend the Second Law of Thermodynamics: energy gradients are exploited by mutually constraining constraints to generate and maintain those very constraints, at the expense of increased entropy production in the environment exterior to closure.

Based on this brief sketch of the OA, we can say how it explains the needful freedom that Jonas identified as the ontology of life as such. Any system that couples thermodynamic openness and organizational closure will exhibit needful freedom precisely to the extent that, as closed, the system is buffered against a range of degrees of kinds of causal fluctuations in the environment and hence is “free”. At the same time this relative independence requires the exploitation of energy gradients, sources of chemical energy, and molecular constituents used in metabolic processes – hence “needful”. The OA allows us to specify in scientific terms what makes the needful freedom of life an ontological surprise – that is, we can specify the precise kind of structural organization that makes the causal dynamics of life distinct from the causal dynamics of physico-chemical systems and hence not derivable from the equations that describe those dynamics. If that is right, however, should we retain Jonas’s dictum that life can only be recognized as such by something that is itself alive? If the reality of teleology is established in scientific explanation, it is less clear what role (if any) there is for phenomenology to play.

The promise of the OA to meet Jonas’s challenge to cybernetics now depends on two further issues: whether the OA succeeds as an account of teleology, and whether the OA should be understood in terms of circular causation. If the OA is best understood as explaining teleology in terms of circular causation, then it can be used as a partial refutation of Jonas’s criticism of cybernetics. Jonas’s critique of cybernetics, recall, was that teleology cannot be explained in terms of circular causation. I shall now argue that the OA succeeds, *pace* Jonas, at explaining teleology in terms of circular causation, but only because the kind of circular causation involved in the OA is very different from the kind of circular causation described by early cybernetics. Put otherwise, the cyberneticists were right to believe that circular causation could explain teleology in naturalistic terms, but mistaken about the precise character of the circular causation required.

The duality of thermodynamic openness and organizational closure should be understood as a duality of two distinct kinds of circular causation, which I shall call ‘synchronic circular causation’ and ‘diachronic circular causation’. In synchronic circular causation, each constraint contributes a causal factor to the totality of constraints which, in turn, exerts a global effect on all of the constraints. There is thus a kind of circular causality to the whole-part relation: each part causally affects the whole and the whole causally affects each part. Organizational closure is therefore a kind of circular causation. In diachronic circular causation, there is a kind of circular causation that emerges from organizational closure with regard to thermodynamic openness. This is because the work done by the constraints exploits energy gradients in order to generate and maintain the constraints that are doing that work. Although neither synchronic nor diachronic circular causation are unique to the organizational approach, the organizational approach theorizes them in distinct terms because of what is unique to the organizational approach: the conception of organizational closure in terms of a closure of constraint and the conceptual distinction between organizational closure and thermodynamic openness as two distinct kinds of causal regime.

If the organizational approach is a theory of circular causation, can we say that it is also a theory of teleology? Mossio and Bich (2017) argue that it is, given a Kantian understanding of teleology as intrinsic purposiveness.[[8]](#footnote-8) The final cause of a self-determining system is itself:

The idea of intrinsic teleology, we submit, does not *merely* point to the realisation of a circular relation between causes and effects but, rather, to the situation in which the activity of a system, by producing some effects, *contributes to specifying the conditions under which the circular relation as such* can occur. It is in this precise sense that the connection between teleology and self-determination is to be understood. (Mossio and Bich 2017, p. 1107; emphasis added)

A self-determining system exhibits intrinsic teleology because of the kind of synchronic and diachronic circular causation that it exhibits: the circular causation of continually contributing to the production of the conditions that allow for the actualization of its own existence. As Mossio and Bich see it, the organizational approach both follows Kant in understanding intrinsic teleology in terms of circular causation but departs from Kant in regarding circular causation can be developed as a formally rigorous and empirically testable theory of empirical science, rather than (as arguably Kant himself did) treating it as a subjective heuristic that had no place in science.[[9]](#footnote-9)

Based on this cursory examination of the organizational approach, we can return to its relevance for Jonas’s criticism of cybernetics. Insofar as Jonas held that the problem with cybernetics is that circular causation could not explain teleology, it would seem that he was mistaken: the organizational approach indicates precisely how circular causation does allow for a naturalistic explanation of teleology. Recall that Jonas objects that cybernetics cannot accommodate the reality of purposes and values because it makes no room for the distinction between a guided missile and a missile that is being piloted by a person; it lacks the Kantian distinction between extrinsic and intrinsic purposiveness. We can see that the OA theorizes the intrinsic purposiveness of organisms in terms of the kinds of causal structures that they posses and that machines, with their extrinsic purposiveness, lack.

The contrast between cybernetics and the organizational approach underscores what is distinct about how the OA explains the needful freedom of the organization in terms of circular causality: not a causality of the transmission of information along a static structure, but a causality of the continual maintenance of the conditions for the existence of the organism. The circular causality is not informational but constitutional. Thus, while the organizational approach shares with cybernetics the emphasis on circular causation or causal loops as the key to naturalizing teleology, it is a very different kind of circular causation. It is for this reason that the OA can succeed where (following Jonas) cybernetics failed.

One advantage of bringing together Jonas and the OA is that doing so allows us to understand and appreciate the insights that led to his criticism of the evolutionary theory of his time, what he simply calls “Darwinism”. Here it is crucial to bear in mind that Jonas’s criticisms are focused upon the Modern Synthesis with its dual dependence on genetic variation and selective pressures. Without disputing the factual basis and logical deductions of the Modern Synthesis, Jonas raises the question as to its ultimate philosophical or metaphysical cogency. The central issue, as he sees it, is the basis of the claim that genetic variation is essentially a matter of “chance” (pp. 44-45). On his view, it is at bottom the need to retain metaphysical materialism that led evolutionary theorists to emphasize the role of chance in the generation of novel phenotypes: anything else would have allowed a divine foot in the door.[[10]](#footnote-10) As Jonas sees it, the central question raised by Darwinism is “whether a mechanistic biology can do justice to the phenomena of life” (p. 52).

This criticism, that Darwinism is too ‘mechanistic’, might seem closer to the pseudo-science of intelligent design than to serious theoretical biology. However, the OA raises a similar criticism about whether evolutionary theory is sufficient as a naturalized theory of teleology. Suppose one were to understand the evolutionary approach to naturalizing teleology as holding that features of biological systems are correctly regarded as functions just in case the present existence of those features is a consequence of the effects of past natural selection on ancestral populations (Millikan 1984). Although it may seem that evolution successfully naturalizes teleology, the OA holds that it fails because of what it presupposes: “evolutionary mechanisms operate *because* they are embodied in the complex organisation of organisms … evolution as an explanatory mechanism actually presupposes the existence of organisms” (Moreno and Mossio 2015, xxi-xxii; emphasis added). Since evolving in response to selective pressures is something that populations of organisms do, evolutionary histories can explain why traits with teleological functions were selected for (due to past natural selection acting on the ancestral populations). But selection does not explain teleological functions in the first place, because evolution presupposes that there are organisms that have any goal-oriented activity at all, relative to which any component processes are intelligible *as* functions.[[11]](#footnote-11)

The importance of this distinction is, I think, underscored by the shift in evolutionary theory from typological to populational thinking (Mayr 1959). As long as pre-Darwinian evolutionary thinking was typological, and species were understood as kinds, ontologically distinct from the particular organisms in which the properties of that kind were instantiated, it was intelligible that evolutionary processes manifested a kind of teleology distinct from that of biological organisms.[[12]](#footnote-12) Ironically, then, the success of the Darwinian revolution in biology -- a shift from typological to populational thinking (Mayr 1959) and a re-categorization of species from kinds to spatio-temporally extended particulars (Ghiselin 1972) – shows why evolutionary theory *cannot* succeed at naturalizing teleology. We must locate naturalized teleology at the level of individual organisms if it exists at all. The organizational approach locates teleology at a fundamentally different place, at a different spatio-temporal scale, than that of evolutionary lineages. This is not to dispute the explanatory power of the Modern Synthesis or of recent suggestions that the modern synthesis be expanded; it is to say only that, of the many diverse features of living things that can be explained in evolutionary terms, the fundamentally teleological organization of living things is not among them.[[13]](#footnote-13)

 Finally, it should be noted the rise of the evolutionary approach to naturalizing teleology – and Jonas’s criticism of Darwinism – emerged at a time when the organism as such was at its nadir in twentieth century biological theorizing. The modern synthesis integrated sub-organismal biology (Mendelian, then molecular genetics) and supra-organismal biology (evolutionary processes), with little left over for a biology that took organisms seriously. Unlike mainstream 20th century evolutionary biology, Jonas retained a central focus on the ontology of organisms – perhaps owing to his having one foot in the tradition of Continental *Naturphilosophie* (see Michelini et al 2018; Michelini 2020). To the extent that contemporary philosophy of biology is rediscovering the central role of the organism, it seems apt to consider the organizational approach to naturalizing teleology as in keeping with the one aspect of Jonas’s bio-philosophy (the emphasis on the organism), even if it takes issue with another (his argument that teleology cannot be naturalized).[[14]](#footnote-14)

3. Assessing the Jonasian Turn in Enactivism

One reason why Jonas’s phenomenological argument for teleological realism merits careful examination is its influence on recent developments in enactivism or enactive cognitive science. The initial proposal for enactivism involved (among other things) the suggestion that cognitive science needed to consider first-person descriptions of experience and not just theorize about cognition from a third-person point of view (Varela, Thompson, and Rosch 1991). While this included Heidegger and Merleau-Ponty at first, Francisco Varela subsequently began reading Hans Jonas and co-authored a famous article that brought Jonas into conversation with autopoiesis and enactivism (Weber and Varela 2002). This led to what Villalobos and Ward (2015a) call the ‘Jonasian turn’ in enactivism, with important contributions by Ezequiel Di Paolo (Di Paolo 2005) and Evan Thompson (2007), among many others. In Thompson’s formulation, Jonas inspires the idea of life-mind continuity: there are signs of mindedness everywhere that there is life.

Following Villalobos and Ward, by “the Jonasian turn” I mean the claim that Jonas’s phenomenological argument for the reality of teleology should be incorporated into enactivism as a research program in cognitive science. Villalobos and Ward criticize the Jonasian turn on the grounds that phenomenology is the wrong sort of approach for grounding teleological realism. As they see it, Jonasian phenomenology leads to “an antiscientific anthropomorphism” (204) because Jonas’s argument for teleological realism begins with what we know to be true about our own case and then applies that human self-knowledge to all of life. Indeed, they make clear that on their view, Jonas himself understood his project as a critique of modern science. But since it is necessary that a scientific research program be as free as possible of anthropomorphic biases, enactivism as a scientific research program should not use Jonasian phenomenology.[[15]](#footnote-15)

This is not to say that Villalobos and Ward are hostile to teleology as such, insofar as they consider it at least a logical possibility that teleology could be naturalized. What they object to is the premodern or Aristotelian positing of “irreducible intrinsic teleology”, defined as “purposes that receive no further explanation in terms of structure, function, or dynamics”. Responding to their critics, they write:

we left the question whether enactivism might, in the future, provide such a further explanation open. But we also claimed that enactivists such as Varela, Di Paolo, and Evan Thompson do not see the providing of such explanations as their task. When autopoiesis or adaptive dynamics are identified with immanent purposiveness by such enactivists, this is not because an explanation of the emergence of natural purposes from non-teleological dynamics has been given. Rather, it has been employed to allow us to understand the relevant dynamics as already imbued with teleology. (Villalobos and Ward 2005b: 230).

In other words, the problem with the Jonasian turn is not that it introduces teleology into enactivism but that it does so *on non-scientific grounds*, where ‘scientific grounds’ would require showing that teleology is not an irreducible feature of life and mind. Rather, we would need an account of teleology consistent with the principles of modern science (which they claim Jonas rejected), and that would require an account of how teleology emerges from non-teleological dynamics.[[16]](#footnote-16)

It is here that the organizational approach can offer the beginning of a response, because the OA conceptualizes intrinsic purposiveness as emerging from a conjunction of closure and openness: the system is organized so as to be self-determining by virtue of extracting from the environment the energy and material necessary to maintain the closure of constraints that constitute its independence from the environment but which at the same time ensure that the organism must be coupled to that environment. By beginning with an account of the causal dynamics that ground attribution of purposiveness and theorizing functions and malfunctions in terms of their contribution to organismal goals, the organizational approach would seem to do exactly what Villalobos and Ward claim enactivists have not done: specify in terms acceptable to modern science how teleology emerges from non-teleological dynamics.[[17]](#footnote-17)

At this point, however, it is also crucial to distinguish between enactivism as a research program in cognitive science and enactivism as a philosophy of nature – a distinction that enactivists have been less than clear on. As Meyer and Brancazio (2022) put it, there is little evidence of a Kuhnian scientific revolution in cognitive science today – even if enactivism offers an alternative to mainstream cognitive neuroscience, there simply has not (yet) been an accumulation of anomalies sufficient to motivate a shift from the cognitivist paradigm to the enactivist paradigm. This is not to say that enactivism is wholly defunct – on the contrary, Meyer and Brancazio are willing to allow that enactivism is viable as a philosophy of nature. A philosophy of nature does not guide first-order theorizing but offers a retrospective comprehensive account of what science has discovered and what it is likely to discover. The task of a philosophy of nature is to offer a comprehensive account in which the results of the sciences are intelligible in ways that would not be taken in isolation.

If the Jonasian turn in enactivism is problematic with regard to enactivism as a scientific research program, it does not follow that the Jonasian turn is problematic with regard to enactivism as a philosophy of nature. As a scientific research program, enactivists should build their project on the organizational approach *rather than* on Jonas’s biological phenomenology. The OA allows enactivists to have what they want from Jonas – an argument for teleological realism – but one that is fully consistent with the rejection of anthropomorphism that modern science demands. A philosophy of nature does not make the same demand; all that it requires is that a coherent picture be offered that illuminates scientific results with a distinct kind of intelligibility. Specifically, enactivism as philosophy of nature articulates a conception of mindedness as holistic, dynamic, non-mechanistic (hence non-computational and non-re-representational). In the recent articulation by Di Paolo, Cuffari, and De Jaegher (2018), enactivism as philosophy of nature is most fundamentally a rejection of hylomorphism (99-101) in favor of an ontology of active matter inspired by Gilbert Simondon. This is, to be sure, a fascinating and provocative way of thinking about what enactivism is all about, but it is a development in enactivism as a philosophy of nature and not (or at least not obviously) a development in enactivism as a research program in cognitive science. Whether Jonas’s biological phenomenology is consistent with enactivism in this sense is an open question, but it cannot be decided as easily as whether the Jonasian turn should be accepted or rejected by enactivism as a paradigm of cognitive science.

4. Conclusion

At the outset, I took an interest in Hans Jonas’s critique of cybernetics because that critique takes places in a remarkable moment in 20th century philosophy: it is the moment when Continental *Naturphilosophie* first encountered the attempt by the architects of the Information Age to definitively resolve the age-old debate between materialism and vitalism. In this encounter, the dialectic of teleology vs mechanism undergoes a profound transformation: it is now no longer a question of whether teleology must be eliminated from a naturalistic worldview, but rather, how adequately teleology can be accommodated by a naturalistic worldview. The cyberneticists took themselves to be resolving the debate by showing precisely how to naturalize teleology; Jonas’s critique demonstrates that they failed. In doing so, however, cybernetics opened up the research program that has led, via Rashevsky and Rosen, Maturana and Varela, to the organizational approach of today. If the OA succeeds where cybernetics fails, it would mean that, *pace* Jonas, the needful freedom of organisms can be explained by circular causation after all. This is not to say that the dialectic of teleology and naturalism has been definitively resolved, but that we have moved onto a new terrain on which the debate shall henceforth be conducted.

**Acknowledgments:** I would like to thank Luca Corti, Andrea Gambarotto, Joseph Jebari, and Auguste Nahas for extensive comments on previous drafts. I am grateful to the two anonymous commentators for their detailed feedback that has vastly improved this paper. Finally, I would also like to thank the guest editors for this special issue of *Topoi* for encouraging the submission of this paper

Works Cited

Corti, L. (2022) The ‘Is’ and the ‘Ought’ of the Animal Organism: Hegel’s Account of Biological Normativity. *History and Philosophy of the Life Sciences* 44 (2):1-22

Coyne, L. (2016) Phenomenology and Teleology: Hans Jonas’s Philosophy of Life. *Environmental Values* 26: 297-315.

Deacon, T. (2012) *Incomplete Nature: How Mind Emerged From Matter.* W. W. Norton.

Di Paolo, E. (2005) Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences* 4(4): 429-452.

Di Paolo, E., H. Cuffari, and H. De Jaegher (2018) *Linguistic Bodies: The Continuity of Life and Language*. The MIT Press.

Dupuy, J. (2009) *On the Origins of Cognitive Science: The Mechanization of the Min*d. The MIT Press, Cambridge.

Frank, L. (1948) Foreword to Special Issue on Teleological Mechanisms. *Annals of New York Academy of Sciences* 50(4):189-196.

Galison, P (1998) The Americanization of unity. *Daedalus* 127(1):45-71.

Gambarotto, A (2020) Teleology, Life and Cognition: Reconsidering Jonas’s Legacy for a Theory of the Organism. *Natural Born Monads* ed. Andrea Altobrando and Pierfranscesco Biasetti. De Gruyter.

Gambarotto, A and A. Nahas (2022) Teleology and the organism: Kant’s controversial legacy for contemporary biology. *Studies in History and Philosophy of Science Part A* 93:47-56.

Garcia-Valdecasas, M. (2021) On the naturalisation of teleology: self-organization, autopoiesis and teleodynamics. *Adaptive Behavior* 30(2): 103-117.

Ghiselin, M. (1972) *The Triumph of the Darwinian Method*. University of California Press, Berkely.

Hverven, S. and T. Netland (2021) Projection or Encounter? Investigating Hans Jonas’s case for natural teleology. *Phenomenology and the Cognitive Sciences*. https://doi.org/10.1007/s11097-021-09748-9

Jonas, H. (2001) *The Phenomenon of Life: Toward a Philosophical Biology*. Northwestern University Press, Evanston

Kee, H. (2021) Phenomenology and naturalism in autopoietic and radical enactivism: exploring sense-making and continuity from the top down. *Synthese* 198: 2323-2343.

Mayr, E. (1959). Darwin and the evolutionary theory in biology. In: Meggers B. J. (Ed.), *Evolution and Anthropology: A Centennial Appraisal* (pp. 1–8). New York: Theo Gaus’ Sons, Inc

Michelini, F., M. Wunsch, and D. Stederoth (2018) Philosophy of nature and organisms’s autuonomy: on Hegel, Plessner, and Jonas’s theories of living beings. *History and Philosophy of the Life Sciences* 40(3):1-27.

Michelini, F. (2020) The Paradox of the Living: Jonas and Schelling on the Organism’s Autonomy. *Rivista di estetica: Rethinking Schelling. Nature, Myth, Realism* 74:139-157.

Millikan, R. (1984) *Language, Thought, and Other Biological Categories*. The MIT Press.

Montévil, M. and M. Mossio (2015) Biological organisation as closure of constraints” *Journal of Theoretical Biology* 372 (May): 179-191.

Moreno, A. and M. Mossio (2015*) Biological Autonomy: A Philosophical and Theoretical Enquiry*. Springer, New York

Moss, L. (2004) *What Genes* Can’t *Do*. The MIT Press, Cambridge.

Mossio, M. and L. Bich (2017) What makes biological organisation teleological? *Synthese* 194(4):1089-114.

Nicholson, D. (2014) The Return of the Organism as a Fundamental Explanatory Concept in Biology. *Philosophy Compass* 9(5):3470359.

Piccinini, G. (2020) *Neurocognitive Mechanisms: Explaining Biological Cognition*. Oxford University Press.

Varela, F., E. Thompson, and E. Rosch (1991) *The Embodied Mind*. The MIT Press

Villalobos, M and D. Ward (2015) Lived Experience and Cognitive Science: Rethinking Enactivism’s Jonasian Turn. *Constructivist Foundations* 11(2): 204-233.

Walsh, D. M. (2015) *Organisms, Agency, and Evolution*. Cambridge University Press, New York.

Weber, A. and F. Varela (2002) Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences* 1: 97-125.

Wiener, N (1948) *Cybernetics: or Control and Communication in the Animal and the Machine*. The MIT Press, Cambridge.

Zammito, J. (2006) Teleology then and now: The question of Kant’s relevance for contemporary controversies over function in biology. *Studies in History and Philosophy of Biology and Biomedical Sciences:* 748-770.

1. Wiener thought the vitalism-mechanism stalemate had been overcome by cybernetics: “Vitalism has won to the extent that even mechanisms correspond to the time-structure of vitalism; but as have seen, this victory is a complete defeat, for from every point of view which has the slightest relation to morality or religion, the new mechanism is fully as mechanistic as the old … the whole mechanist-vitalist controversy has been relegated to the limbo of badly posed questions” (Wiener 1948, p. 44); cf. “Information is information, not matter or energy. No materialism which does not admit this can survive at the present day” (ibid., p. 132). In other words, cybernetics is a new kind of mechanisms that accommodates what vitalism had attempted to explain: the reality of self-organizing systems and the emergence of teleology. [↑](#footnote-ref-1)
2. “Cybernetics: A Critique” first published in 1952, all page numbers refer to the reprinted edition in *Phenomenon of Life* (2001). [↑](#footnote-ref-2)
3. For recent explications of Jonas’s bio-philosophy, see Coyne (2017), Pommier (2017), and Hverven and Netland (2021). [↑](#footnote-ref-3)
4. This section is primarily based upon Moreno and Mossio (2015), Montévil and Mossio (2015), and Mossio and Bich (2017). For reasons of space, I omit from consideration how the OA builds upon but differs from the closely related autopoiesis approach developed by Maturana and Varela. A comprehensive discussion of the OA would also need to discuss its deep roots in Aristotle, Kant, and Hegel. [↑](#footnote-ref-4)
5. Amongst their predecessors they cite not only Kant but also Rashevsky, Rosen, Kauffman, Pattee, Juarrero, Ganti, and Piaget. [↑](#footnote-ref-5)
6. It would be a further question as to why exactly, on the organizational approach, distinguishes self-determining systems (organisms) and self-organizing systems (autocatalytic networks, dissipative structures), since it is not entirely clear how adding additional constraints is sufficient explanation. [↑](#footnote-ref-6)
7. [↑](#footnote-ref-7)
8. For a recent defense of the Kantian credentials behind the organizational approach with regard to teleology, see Gambarotto and Nahas (2022); for a recent treatment of Hegel’s philosophy of nature and the OA, see Corti (2022). [↑](#footnote-ref-8)
9. I understand that this claim is contentious, but see Zammito (2006). [↑](#footnote-ref-9)
10. This is not to deny that there is no empirically detectable mechanism that can foresee what traits will be adaptive when the environment changes and then cause those traits to come; it is to question whether the supposition of metaphysical materialism, specifically Epicureanism, influenced the attribution of more importance to chance than what was required by the ontological commitments of empirical theory. [↑](#footnote-ref-10)
11. For a comprehensive assessment of why selection must assume teleology and cannot explain it, see Walsh (2015); see also Piccinini (2020), pp. 71-81. [↑](#footnote-ref-11)
12. See Moss (2004) for the shift from ontogenetic teleology to phylogenetic teleology correlated with the rise of mechanistic science. [↑](#footnote-ref-12)
13. See Mossio and Bich (2017) for a more detailed comparison of the organizational and evolutionary approaches to naturalizing teleology. [↑](#footnote-ref-13)
14. For recent work on the return of the organism in philosophy of biology, see Nicholson (2014) and Walsh (2015). [↑](#footnote-ref-14)
15. For a defense of Jonas in response to Villalobos and Ward, see Hverven and Ward (2021); for an argument that enactivists should reject Jonas’s phenomenology but still use Merleau-Ponty, see Kee (2021). [↑](#footnote-ref-15)
16. For a similar assessment of the contemporary significance of Jonas’s bio-philosophy, see Gambarotto (2020). [↑](#footnote-ref-16)
17. Arguably the organizational approach at best specifies the kind of nonlinear or circular causality that grounds teleology, but not how teleology itself emerges from non-teleological systems. It is here that I think Deacon’s (2012) distinctions between homeodynamics, morphodynamics, and teleodynamics should be considered an advance upon the organizational approach. See Garcia-Valdecasas (2021) for a contrast between the OA and teleodynamics. [↑](#footnote-ref-17)