

## The British Society for the Philosophy of Science

---

An Organizational Account of Biological Functions

Author(s): Matteo Mossio, Cristian Saborido and Alvaro Moreno

Source: *The British Journal for the Philosophy of Science*, Vol. 60, No. 4 (Dec., 2009), pp. 813-841

Published by: Oxford University Press on behalf of The British Society for the Philosophy of Science

Stable URL: <https://www.jstor.org/stable/25592036>

Accessed: 11-02-2019 21:29 UTC

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

*Oxford University Press, The British Society for the Philosophy of Science* are collaborating with JSTOR to digitize, preserve and extend access to *The British Journal for the Philosophy of Science*

# An Organizational Account of Biological Functions

Matteo Mossio, Cristian Saborido, and  
Alvaro Moreno

---

## ABSTRACT

In this paper, we develop an organizational account that defines biological functions as causal relations subject to closure in living systems, interpreted as the most typical example of organizationally closed and differentiated self-maintaining systems. We argue that this account adequately grounds the teleological and normative dimensions of functions in the current organization of a system, insofar as it provides an explanation for the existence of the function bearer and, at the same time, identifies in a non-arbitrary way the norms that functions are supposed to obey. Accordingly, we suggest that the organizational account combines the etiological and dispositional perspectives in an integrated theoretical framework.

- 1 *Introduction*
  - 2 *Dispositional Approaches*
  - 3 *Etiological Theories*
  - 4 *Biological Self-maintenance*
    - 4.1 *Closure, teleology, and normativity*
    - 4.2 *Organizational differentiation*
  - 5 *Functions*
    - 5.1 *C<sub>1</sub>: Contributing to the maintenance of the organization*
    - 5.2 *C<sub>2</sub>: Producing the functional trait*
  - 6 *Implications and Objections*
    - 6.1 *Functional versus useful*
    - 6.2 *Dysfunctions, side effects, and accidental contributions*
    - 6.3 *Proper functions and selected effects*
    - 6.4 *Reproduction*
    - 6.5 *Relation with other 'unitarian' approaches*
  - 7 *Conclusions*
-

## 1 Introduction

During the last 40 years, the concept of function has been investigated by an increasingly rich literature in the field of philosophy of science.<sup>1</sup> The main reason for this interest lies in the fact that functions seem to call for the naturalization of their teleological and normative dimensions.

First, functions seem to have a genuine explanatory role in accounting for the *existence* of function bearers. Affirming that—to cite a classical example—‘the function of the heart is to pump blood’ is to say that this effect has specific relevance in explaining the existence, structure, and morphology of hearts (see also Buller [1999], pp. 1–7). Functional attributions thus introduce a teleological dimension in the structure of explanation, insofar as the existence of a trait could be explained by appealing to some specific effects or consequences of its own activity. In this respect, the explanatory power of functions has to be reconciled with their teleological ‘flavour’.

Second, the concept of function possesses a normative dimension to the extent that it refers to some effect that the trait is *supposed* to produce (Hardcastle [2002], p. 144). As McLaughlin ([2001]) has pointed out, functions are means–end relationships, which go beyond standard efficient causation. Attributing functions to traits implies a reference to some specific effect, which constitutes a criterion against which the activity of the trait can be normatively evaluated. The normative dimension of functions requires, then, a theoretical justification of the criteria upon which functional relations are identified among all causal relations.

The philosophical debate has addressed both issues, and several accounts have been proposed to interpret the concept of function, in order to make it compatible with the accepted structure of scientific explanation. Broadly speaking, the existing philosophical accounts of functions can be classified into two main traditions, which mainly diverge on the explanatory strategies adopted to deal with the problem of teleology.

A first tradition, usually labelled ‘etiological’, has tried to justify and naturalize the teleological dimension of functions by appealing to a scientifically acceptable causal explanation. In the mainstream formulation, etiological approaches appeal to a historical-selective causal process, through which the existence of current functional traits is the consequence of the selection exerted on the effects of previous occurrences of the trait. A second tradition, called ‘systemic’ or ‘dispositional’, discards the teleological dimension of functional attributions as a relevant explanandum by interpreting functions as causal means–end relations at work in a system. From this second perspective,

<sup>1</sup> Several collections surveying the recent philosophical debate have appeared: see (Ariew et al. [2002]; Buller [1999]; Allen et al. [1998]).

functions do not explain the existence of the bearer; they refer to current contributions of functional traits to some capacity of the system to which they belong. Both the etiological and the dispositional traditions have addressed the issue of normativity. As we will discuss, the etiological approach usually interprets norms in terms of the evolutionary conditions of existence of the functional trait. In contrast, a debate has arisen within the dispositional tradition between authors claiming that an adequate theory of functions has to ground their normative dimension, and authors for whom not even a grounding of norms is required.

The two traditions have been widely discussed and their qualities and limits precisely analyzed, along with their mutual relations. Most authors seem to agree that etiological and dispositional approaches provide alternative definitions of functions, to the extent that the explanation of the existence of the functional trait seems to be conceptually independent from the explanation of its contribution to a capacity of the current system, and vice versa. Some authors have proposed a pluralistic solution to the problem of alternative definitions of functions, arguing that the etiological and dispositional approaches actually provide two complementary concepts of function (Allen and Bekoff [1995]; Godfrey-Smith [1994]; Millikan [1989]). Other authors, such as Kitcher ([1993]), Walsh ([1996]), and Walsh and Ariew ([1996]), have claimed that there is, in fact, a unique concept of function, in which the etiological and dispositional formulations can be subsumed as special cases.<sup>2</sup>

In this paper, we develop an *organizational* account of biological functions that shares with previous accounts the notion that the etiological and dispositional perspectives can be integrated into a unique conceptual framework. In particular, we will interpret functions as causal relationships at work in the organization of biological systems, which at the same time (i) explain the existence of the function bearer and (ii) constitute means–end relationships contributing to some distinctive capacity of the current system being analyzed. The gist of the argument will be that functions are inherently related to the idea of a *closed and differentiated self-maintaining organization*, which provides an adequate grounding for their teleological and normative dimensions. A relevant implication of the organizational account is that it is likely to provide a unified account for functional attributions to both biological traits and artefacts. In this paper, however, we will restrict the analysis to the functions of biological traits, and leave the case of artefacts—which will require a detailed examination—for a future enquiry.

The structure of the paper is as follows. In Sections 2 and 3, we present a critical analysis of the existing dispositional and etiological approaches and

<sup>2</sup> In Section 6.5, we will discuss in more detail these ‘unitarian’ approaches, as Buller ([1999]) labels them.

discuss their relative strengths and weaknesses. Our analysis indicates the importance of developing an account of biological functions that combines the virtues of these two approaches and can ground both the teleological and normative dimensions of functions in the *current* organization of the system. In Section 4, we introduce a description of biological organization in terms of self-maintenance, by arguing that two of its specific properties, closure and differentiation, adequately naturalize teleology and normativity, and then ground functional attributions in the current organization of the system. In Section 5, we articulate our characterization of biological functions from the organizational perspective. In Section 6 we discuss some consequences and properties of our proposal, as well as some possible objections and counter-examples. Finally, in Section 7, we conclude by discussing the relation between our proposal and similar ones.

## 2 Dispositional Approaches

In the philosophical debate on functions, several authors have developed accounts of functions based on the idea that functional attributions refer to *current* features of the system under examination. Following the original proposal by Nagel ([1977], [1961]), these accounts discard teleology as a constitutive dimension of the concept of function. Thus, functions do not refer to a causal process that would explain the existence of the function bearer by appealing to its effects. Rather, functional relations are interpreted as a particular class of causal effects or dispositions of a trait, means–end relationships contributing to some distinctive capacity of the system to which it belongs. On the basis of this common theoretical stance, these approaches have been labelled ‘causal role’, ‘dispositional’ or ‘forward-looking’, as opposed to ‘backward-looking’ etiological ones. In this paper, we will use the general label ‘dispositional’ to refer to this class of theories.

The philosophical agenda of dispositional approaches focuses on providing naturalized (i.e., grounded in some constitutive features of the system and not related to an extrinsic evaluative decision of the observer) and appropriate (i.e., in accordance with both scientific and everyday usage) criteria to identify what counts as a target capacity of a functional relationship, from which the relevant norms could be deduced, and the different dispositional approaches have proposed various criteria to identify these target capacities.

The more classical dispositional approach is the ‘systemic approach’ (SA), which defines a function  $F$  as the contribution of a process  $P$  to a distinctive higher-level capacity  $C$  of the system  $S$  to which it belongs (Craver [2001]; Cummins [1975]; Davies [2001]). In the SA, explaining functions means analyzing a given higher-level capacity of the system into the capacities of the system’s components, which jointly concur to the emergence of the higher-level capacity.

Accordingly, as Davies clarifies, the SA restricts functional attributions to *hierarchically organized* systems (Davies [2001], p. 85; see also Craver [2001], Section 4). The SA extends functional ascriptions to relations at work in a wide range of systems and capacities and dissolves the problem of teleology of functions by reducing them to any causal contribution to a higher-level capacity. In turn, the normative dimension of functions is reduced to the constraint that the causal effect must contribute to a higher-level capacity: in particular, no reference to a ‘benefit’ for the system is required when identifying functional effects.

Despite its restriction to hierarchically organized systems, the explanatory strategy adopted by the SA is exposed to a major criticism, namely, that it seriously *under-specifies* functional ascriptions, which in turn generates several problems (see also Wouters [2005]). First, the SA fails to draw a principled demarcation between systems whose parts appear to have functions and systems whose parts do not (Bigelow and Pargetter [1987]; Millikan [1989]).<sup>3</sup> Second, the SA lacks a principled criterion to identify the relevant set of contributions for which functional analysis makes sense. Consequently, the SA is unable to adequately distinguish between functional contributions and dysfunctional or irrelevant effects (Millikan [1989]; Neander [1991]). Third, the SA is unable to draw an appropriate distinction between ‘proper’ functions and accidental, useful contributions (Millikan [2002], [1993]).

Because of these fundamental weaknesses of the SA, the ‘goal contribution approach’ (GCA) has attempted to introduce more specific constraints on what makes causal relations properly functional, by linking the concept of function to the idea of *goal-directedness*. In particular, the GCA restricts functional attributions to causal contributions to those (higher-level) capacities constituting the ‘goal states’ of the system (Adams [1979]; Boorse [2002], [1976]). To identify the goal states of a system following a naturalized and non-arbitrary criterion, the GCA has adopted a cybernetic characterization of goal-directedness, based on ideas previously proposed by Rosenblueth *et al.* ([1943]) and developed by Sommerhoff ([1950], [1959]) and other authors. In particular, biological systems can be described as having as their essential goal to survive (and reproduce). Hence, biological functions are dispositions that contribute to these goals.

The explanatory strategy of the GCA consists of restricting functions to the contributions to those higher-level capacities that can be described in a naturalized way as ‘goals’ in cybernetic terms. The main virtue of this perspective

<sup>3</sup> As Craver points out, Cummins’ conditions for performing a functional analysis certainly ‘point to important *symptoms*’ that the considered system is a relevant one (Craver [2001], p. 57). Nevertheless, they are too generic to rule out counter-examples, since many physical dissipative systems, such as hurricanes, in which the activity of microscopic elements causes the emergence of a higher-level pattern (Nicolis [1989]), seem to fit the conditions for being considered as ‘hierarchically organized’, even if we do *not* attribute functions to their parts.

is that it provides an interpretation of functions that, in contrast to the SA, recognizes and substantiates their specificity as means–end causal relationships. Nevertheless, the very choice of interpreting the target capacity in cybernetic terms constitutes the fundamental drawback of the GCA. The cybernetic characterization of a goal-directed system introduces norms whose application is, in fact, not restricted to the relevant kinds of systems and capacities. As Bedau ([1992]) points out, the cybernetic characterization of the goal state is unable to adequately capture the frontier between ‘genuinely’ goal-directed systems (supposedly: biological systems and artefacts) and physical equilibrium systems, which tend to some steady state or state of equilibrium (see also Nissen [1980]). Moreover, as Bedau ([1992]) and Melander ([1997]) argue, cybernetic criteria may interpret dysfunctional behaviours of goal-directed systems as functional. Finally, the GCA account, much like the SA account, lacks the theoretical resources to distinguish between functions and accidental contributions to a goal state. In sum, the GCA still seems to under-specify functional attributions, and in some cases it appears even to be a less satisfactory account than the SA.

Advocates of the SA and of the GCA often reply by explicitly denying that a theory of functions should provide principled criteria for the kind of distinctions invoked in criticisms. For instance, Craver argues that deciding what is the ‘right’ capacity the item is supposed to contribute to depends on ‘background assumptions’ that are external to the theory (Craver [2001], p. 71). Davies ([2001]) develops an argument against the need to distinguish theoretically between proper functions and accidental contributions. Similarly, Boorse argues that the distinction between ‘functions performed’ and ‘functions possessed’ is inevitably vague, and depends on ‘how often or consistently the function gets performed’ (Boorse [2002]). This position, however, appears to be somewhat at odds with the way in which both science and common language ascribe functions and is therefore, to some extent at least, unsatisfactory.

The third main dispositional perspective proposes to identify functions with causal contributions of components to the life chances (or fitness) of the system (Bigelow and Pargetter [1987]; Canfield [1964]; Ruse [1971]). Here we will focus on Bigelow and Pargetter’s account, which is probably the best known and most widely discussed.

Bigelow and Pargetter have proposed the ‘propensity view’, according to which ‘something has a (biological) function just when it confers a survival-enhancing propensity on the creature that possesses it’ (Bigelow and Pargetter [1987], p. 108). By appealing to survival in terms of enhancing propensities as the goal of a functional relation, the propensity view succeeds in restricting functions to components of biological entities. Moreover, Bigelow and Pargetter’s reference to survival-enhancing *propensity* is intended to avoid functional attributions to contingent and/or accidental contributions to survival, which would be contrary to intuition and common use. As the authors clarify: ‘What

confers the status of a function is not the sheer fact of survival-due-to-a-character, but rather, survival due to the propensities the character bestows upon the creature' (Bigelow and Pargetter [1987], p. 107).

Yet, as McLaughlin perceptively argues (McLaughlin [2001], pp. 125–8), the appeal to propensities does not fully succeed in restricting functional attributions to the relevant cases. Even by restricting propensity to the current environment (the 'natural habitat', in Bigelow and Pargetter's terms), it is possible to imagine, for each specific effect produced by a trait, a situation in which that specific effect would confer a (possibly low) propensity that enhances survival, and thus have a function. The problem is that propensities to enhance survival in virtual (but not impossible) situations correspond, in a forward-looking approach, to *actual* functions of the existing trait. Moreover, to the extent that the specific contribution of the trait would presumably change according to the particular condition, each trait in fact possesses an indefinite list of actual functions. Again, the propensity view fails to provide an adequately restricted definition of what counts as a functional relation. All (biological) functions are survival-enhancing contributions, but not all survival-enhancing contributions are functions. Appealing to propensities does not solve the problem.

To summarize, the main virtue of the dispositional approaches is their capacity to capture a fundamental dimension of the concept of function, i.e., the fact that it points to something more than mere causal relations: functions refer to current means–end relationships, and more specifically to current contributions of components to the emergence of a *target* capacity of the containing system. Yet, dispositional approaches in the end fail to provide a fully satisfactory grounding for the normativity of functional attributions. Dispositional definitions turn out to be systematically under-specified, allowing functional ascriptions to classes of systems and/or capacities for which the functional discourse does not seem to be usual or relevant. In a word, the price paid for excluding the teleological dimension as a proper explanandum is not fully compensated by a satisfactory foundation of the normative dimension.

In fact, most of the existing literature has favoured a different approach, according to which an adequate understanding of functional attributions has to deal with the problem of teleology. In particular, both the teleological and normative dimensions are conceived as being inherently related to the *etiology* of the functional trait.

### 3 Etiological Theories

The mainstream philosophical theory of functions is the etiological approach (Wright [1973], [1976]; Millikan [1984], [1989]; Neander [unpublished], [1991]; Godfrey-Smith [1994]). The etiological approach defines a trait's function in terms of its etiology (i.e., its causal history): the functions of a trait are past



effects of that trait that causally explain its current presence. The etiological approach explicitly takes the issue of teleology as the central problem of a theory of functions. In this sense, this approach represents the heritage of Hempel's formulation of the problem (Hempel [1959]). In contrast with Hempel's own point of view, however, the advocates of an etiological approach have tried to justify the idea that functional attributions explain the existence of the function bearer in a scientifically valid way.

Historically, the first etiological approach was proposed by Wright, who defined functions as follows.

The function of X is Z means:

- (1) X is there because it does Z
- (2) Z is a consequence (or result) of X's being there (Wright [1976], p. 48).

Wright's definition explicitly appeals to a form of causal 'loop', in which the effect of a trait helps to explain its existence. The scientific validity of Wright's definition has been questioned and, moreover, several obvious counter-examples have been formulated (see, for instance, Boorse [1976]).

In order to ground the teleological dimension of functions without adopting an unacceptable interpretation of the causal loop described by Wright, mainstream etiological accounts, usually called 'selected effect (SE) theories', have appealed to selection as the causal process, which would adequately explain the existence of the function bearer by referring to its effects. In fact, according to the SE theories, functional processes are not produced by the same tokens of which they are supposed to explain the existence. Instead, the function of a trait is to produce the effects for which past occurrences of that trait were selected by natural selection (Godfrey-Smith [1994]; Millikan [1989]; Neander [1991]). Selection explains the existence of the *current* functional trait because the effect of the activity of *previous* occurrences of the trait gave the bearer a selective advantage. The main consequence of this explanatory line is its historical stance: what makes a process functional is not the fact that it contributes in some way to a present capacity of the system, but that it has the right sort of selective history.

By interpreting functions as selected effects, the SE theories are able not only to deal with the problem of teleology, but also to ground the normativity of functions. By defining functions as effects subject to an evolutionary causal loop, the SE theories identify the norms of functions with their *evolutionary conditions of existence*: the function of a trait is to produce a given effect because, *otherwise*, the trait would not have been selected, and therefore would not exist.

Several virtues of the SE theories are often emphasized, such as their capacity to exclude functional attributions to traits of physical systems, to unambiguously identify functions among the whole set of all processes occurring in a

system and, in particular, to draw a frontier between functions and accidental useful effects. In addition, as their advocates often emphasize, the SE theories offer a suitable justification of malfunction: a trait malfunctions if it fails to do what it was selected for (Millikan [1989]; Neander [1991]).

Yet, the SE theories have their own weaknesses, which have been extensively discussed in the literature (see, for instance, Boorse [1976]; Cummins [2002]; Davies [2000], [1994]). We will focus here on a specific one, which Christensen and Bickhard ([2002]) have labelled *epiphenomenalism*. The crucial drawback of the SE theories' explanatory line is the implication that functional attributions have no relation to the *current* contribution of the trait to the system, since they point solely to the selective history of the trait. This implication is problematic because it is at odds with the fact that functional attributions to biological structures do seem to have some relation—captured by the dispositional approaches—to what they currently do, and not only to what explains their current existence. Accordingly, the SE theories provide an account that is problematically epiphenomenal, in the sense that it maintains that the attribution of a function does not provide information about the 'phenomenon' (the current system) being observed. From the perspective of the SE theories, a function does not describe anything about the current organization of the system being analyzed.<sup>4</sup>

Advocates of the SE theories are of course well aware of this problematic implication (Millikan [1989]; Neander [1991]). The main argument proposed to defend the etiological viewpoint consists of pointing out that the idea that functional attributions refer to current features of the system, even if appealing, is in fact irrelevant, since it cannot be adequately grounded, as the limits of the dispositional approaches seem to suggest. Dispositional interpretations may possibly provide a general characterization of 'useful' contributions, without being able to distinguish between proper functions (*functions of*) and accidental or secondary contributions (*functions as*). Only the evolutionary etiology can ground the right sort of normativity required to identify the function of a trait.

#### 4 Biological Self-maintenance

The outcome of the brief critical survey proposed in Sections 2 and 3 is that current theories of functions seem to face a dilemma, arising from the way in

<sup>4</sup> To solve some difficulties inherent in the previous formulations of the etiological theories (mainly that they attribute proper functions to effects that are, in fact, no longer functional in the current system), Godfrey-Smith ([1994]) has proposed a 'modern history theory' of function. In his approach, functions are 'dispositions or effects a trait has which explain the recent maintenance of the trait under natural selection' (Godfrey-Smith [1994], p. 199). While it successfully counters several objections raised against previous versions of the theory, Godfrey-Smith's account is no better placed to deal with the problem of epiphenomenalism. More precisely, as McLaughlin ([2001], p. 116) points out, Godfrey-Smith's account (which is explicitly a historical one), by reducing the cases in which it attributes functions to currently non-functional traits, possibly reduces 'uncooperative cases', but it does not provide a principled solution to the problem.

which they deal with the two main issues related to the concept of function, i.e., its teleology and its normativity. Dispositional theories try to account for functions in terms of current contributions to some target capacity of a system, and discard the teleological dimension, but seem unable to provide fully adequate normative criteria on functional attributions. Etiological theories, on the other hand, try to account for both the teleological and normative dimensions of functions, but are unable to justify how functional attributions may refer to features and properties of the current system being analyzed.

In this section, we will develop an organizational account (OA) of functions, which tries to overcome the dilemma. The core of the OA is the idea that functional attributions do account *at the same time* for the existence of functional traits and for their current contribution to a systemic capacity, since functions make sense only in relation to a specific kind of physico-chemical organization, which is characteristically at work in biological systems.

To attain this objective, the OA relies on a scientific and conceptual framework, developed during the last 40 years in the domains of theoretical biology, complex systems theory, and far-from-equilibrium thermodynamics, which provides an understanding of the constitutive organization of biological systems in terms of *self-maintenance*. Until very recently, no philosophical account of functions had explicitly tried to establish a connection with these theories, possibly because they were being developed in parallel with the philosophical debate on functions and could not provide a reliable scientific basis. Today, these theories have been considerably developed, and we will suggest that they offer interesting new conceptual tools to investigate the nature of functions.

The specific aim of this section is to describe biological systems as a specific class of self-maintaining systems, by focusing on two of their fundamental properties, crucially involved in the grounding of functional attributions: organizational *closure* and organizational *differentiation*. We will argue that biological systems, as organizationally closed and differentiated self-maintaining systems, possess the necessary properties for adequately grounding both the teleological and normative dimensions of functions in their current organization.

#### 4.1 Closure, teleology, and normativity

The OA relies on the understanding of biological systems as a sophisticated and highly complex example of natural *self-maintaining systems*.

In the last 40 years, the study of self-maintaining systems has become a thriving area of investigation in many scientific disciplines, including physics, chemistry, and biology. Self-maintenance is a widespread phenomenon in nature, exhibiting different degrees of complexity. Its minimal expression consists of the so-called ‘dissipative structures’ (Glansdorff and Prigogine [1971]; Nicolis and Prigogine [1977]; Nicolis [1989]). Dissipative structures are systems

in which a huge number of microscopic elements adopt a global, macroscopic ordered pattern (a 'structure') in the presence of a specific flow of energy and matter in far-from-thermodynamic equilibrium (FFE) conditions. In turn, the macroscopic pattern exerts at least one 'top-down' constraint that contributes to the maintenance of the FFE flow of energy and matter enabling the persistence of the microscopic dynamics (Ruiz-Mirazo [2001], p. 59). Insofar as they make an active contribution to maintaining the FFE conditions required for their own existence, dissipative structures can be described as self-maintaining systems.

A standard and relatively simple example of a dissipative structure are the 'Bénard cells', i.e., macroscopic structures that appear spontaneously in a liquid when heat is applied from below (Chandrasekhar [1961]). In the initial situation, in which there is no difference in temperature between upper and lower layers, the liquid tends to be in a state of thermodynamical equilibrium, and appears perfectly uniform (in terms of the statistical distribution of the kinetic energy of the molecules). When heat is applied from below, the liquid is traversed by a continuous flow of energy that modifies its molecular structure and generates a linear variation in temperature, density, and pressure between the lower and upper planes. If the temperature in the bottom layer is increased up to a specific threshold, the liquid's dynamics changes dramatically: in *far* from thermodynamic equilibrium, the random movements of the microscopic molecules spontaneously become ordered, creating a macroscopic pattern, usually called 'convection cells'. In each cell, billions of microscopic molecules rotate in a coherent manner along a hexagonal path, each cell rotating either clockwise or anticlockwise, and always in the opposite direction from that of its immediate neighbours in a horizontal plane.

The crucial property of Bénard cells is that their existence depends not only on a set of boundary conditions (the heat applied from below), but also on the mutual causal relation existing between the macroscopic pattern and the microscopic dynamics. On the one hand, the coherent rotation of billions of molecules generates each macroscopic cell, as we have described. On the other hand, the cells exert a constraining action on the microscopic dynamics, since the fact of belonging to a given cell determines whether a molecule rotates in a clockwise or counter-clockwise direction. Since Bénard cells exist only in the presence of (i) a flow of energy traversing the system (induced by the external heat supply) and (ii) the mutual causal dependence between the macroscopic pattern and the microscopic dynamics, they are an example of self-maintaining systems.

A second example of a dissipative structure is the flame of a candle. In a candle flame the microscopic reactions of combustion generate a macroscopic pattern (the flame itself) that contributes to maintaining the conditions for its own existence. Specifically, the flame makes a crucial contribution to maintaining the microscopic chemical reactions by keeping temperature above the

combustion threshold, vapourizing wax, and inducing convection (which pulls in oxygen and removes combustion products). In turn, the chemical reactions keep generating the combustion and thus the flame itself. As with Bénard cells, the flame works to maintain the conditions that enable it to work (Bickhard [1993], [2000]).

The central conceptual point for the present discussion is that self-maintaining systems can be said to instantiate what we call *organizational closure*, i.e., a circular causal relation between some macroscopic (or higher-level) pattern or structure and the microscopic (or lower-level) dynamics and reactions. Because of the organizational closure, the activity of the system becomes a necessary (even if, of course, not sufficient) condition for the system itself.<sup>5</sup>

In nature, a very broad set of physical and chemical systems (in addition to the example described above, one may mention whirlwinds, hurricanes, oscillatory chemical reactions,<sup>6</sup> and autocatalytic chemical networks) can be pertinently described as self-maintaining systems (Nicolis and Prigogine [1977]; Field *et al.* [1972]; Field and Noyes [1974]). In particular, insofar as they can be described as being subject to organizational closure, biological systems can also be considered as a paradigmatic example of extremely complex self-maintaining systems. Indeed, biological systems contribute in many ways to maintaining a constant exchange of energy and matter with the external environment, which in turn is required to maintain their constitutive organization (their metabolism). As simpler physico-chemical self-maintaining systems, biological systems work to maintain the FFE conditions enabling them to work.

A first important implication of organizational closure is that it provides a criterion to determine in a non-arbitrary way what ‘the goals’ of the system are. The goal states of self-maintaining systems are not such simply because they are characteristic or interesting capacities (as the systemic accounts suggest), or because the systems compensate for deviations from them (as classical cybernetic accounts claim). More fundamentally, in an organizationally closed system the goal states are the stability points (or set of points) through which the system can exist (Barandiaran and Moreno [2008]).

We can now formulate our first thesis. We claim that organizational closure grounds teleology and normativity in a naturalized way. As we have just discussed, the central feature of self-maintaining systems is that (at least some

<sup>5</sup> An obvious but crucial remark. Self-maintenance has nothing to do with ‘independence’ from the environment or external conditions. A system can be said to be self-maintaining *with respect to* a given environment and context, in which it can exert a constraint on some boundary conditions. In this sense, self-maintenance is a context-dependent determination. This point will become important in the following sections.

<sup>6</sup> Such as the well-known ‘Brusselator’ (Nicolis and Prigogine [1977]) and the Belussov–Zabotinsky reaction; other classical examples of oscillatory chemical reactions are the ‘FKN’ model (Field *et al.* [1972]) and the ‘Oregonator’ (Field and Noyes [1974]).

of) their constitutive processes generate some constraints that contribute to maintaining the FFE conditions, which in turn enable the constitutive processes to occur. Hence, organizational closure justifies explaining the existence of a process by referring to its effects: a process is subject to closure in a self-maintaining system when it contributes to the maintenance of some of the conditions required for its own existence. In this sense, organizational closure provides a naturalized grounding for a teleological dimension: to the question ‘Why does X exist in that class of systems?’, it is legitimate to answer ‘Because it does Y’. In the case of the flame, for instance, the existence of the reactions of combustion is explained (at least in part) by taking into account their consequences and, in particular, the generation of a macroscopic flow of energy (the flame), which in turn maintains the conditions in which the chemical reactions of combustion may occur.

Similarly, organizational closure grounds normativity. Because of the organizational closure, the activity of the system has an intrinsic relevance for the system itself, to the extent that its very existence *depends on* the effects of its own activity. Such intrinsic relevance, we hold, generates a naturalized criterion to determine what the system is ‘supposed’ to do. In fact, the whole system (and its constitutive processes) ‘must’ behave in a specific way, *otherwise* it would cease to exist. Accordingly, the activity of the system becomes its own norm or, more precisely, the conditions of the existence of its constitutive processes and organization are the norms of its own activity.

Organizational closure provides, then, a naturalized grounding for both teleology and normativity in the *current* organization of a self-maintaining system. Their contribution to the maintenance of the whole system constitutes at the same time (one of) the *conditions* for the existence and the *norm* of all processes subject to closure. Yet, although necessary in order to naturalize teleology and normativity, organizational closure is not a sufficient condition for functional attributions. Minimal self-maintaining systems, such as flames or hurricanes, do instantiate closure, but their components do not have functions. To have functions, self-maintaining systems must belong to a specific class in which different contributions to the self-maintenance of the system can be distinguished.

## 4.2 Organizational differentiation

Dissipative systems instantiate a *minimal* form of self-maintenance in that they generate a *single* macroscopic pattern constraining their constitutive processes in an organizationally closed way. Dissipative systems make a unique contribution to their own maintenance. Instead, and this is our second main thesis, functional attributions apply to systems in which it is possible to distinguish between different contributions to self-maintenance made by the constitutive

organization. Functional ascriptions require not only closure, but also organizational differentiation.

Organizational differentiation has to be carefully distinguished from what could be labelled the 'material complexity' of a system, i.e., the variety of its internal components. Self-maintaining systems may indeed differ considerably with respect to their material complexity. Whereas many physical dissipative systems possess a rather homogeneous structure in terms of the variety of molecules composing them (e.g., whirlwinds and Bénard cells), other systems, including chemical dissipative systems such as candle flames, have many different molecular components. Certain types of chemical systems (autocatalytic networks, for instance) may even possess a high degree of material complexity.

Even high material complexity, however, does not imply organizational differentiation, and thus does not imply functions. In the case of the flame, for instance, the different chemical components all 'converge' to generate a single macroscopic pattern (the flame), which in turn constrains the microscopic dynamics by maintaining FFE conditions, as already discussed. Accordingly, it is not possible in this case to distinguish between the different ways in which the various components contribute to the self-maintenance of the system. Hence, functional attributions to components of the flame, as well as to all dissipative structures, are not meaningful. It is for this reason, we argue, that we do not attribute functions to simple forms of self-maintaining systems, whatever their material complexity may be.<sup>7</sup>

Organizational differentiation is the second necessary ingredient for functional attributions, the other being closure. Organizational differentiation implies not only that different material components are recruited and constrained to contribute to self-maintenance but, in addition, also that the system itself *generates* distinct structures contributing in a different way to self-maintenance. In other words, material components become candidates for functional attributions only if they have been generated, and are maintained, within and by the organization of the system. A self-maintaining system is organizationally differentiated if it produces different and localizable patterns or structures, each making a specific contribution to the conditions of existence of the whole organization.<sup>8</sup>

<sup>7</sup> The same argument applies to Michael Ruse's example of the water cycle (Ruse [2003], pp. 263–4), used to counter the idea that circular causality suffices to ground functions. The system including the water cycle may possibly be considered a self-maintaining system, but not an organizationally differentiated one. Accordingly, it does not seem to constitute a counter-example to our account.

<sup>8</sup> For a more detailed understanding of the transition between material complexity and functional differentiation, let us take the example of an autocatalytic chemical set. An autocatalytic set is a collection of molecular entities (catalysts), each of which can be created catalytically by other catalysts within the set, such that, as a whole, the set is able to catalyze its own production. In this sense, the set as a whole is said to be 'autocatalytic' (Kauffman [2000]). Consider now the very minimal conceivable autocatalytic set, constituted by a single (though possibly chemically very

The most prominent examples of organizationally closed and differentiated self-maintaining systems are living organisms and in particular, their metabolism. Metabolic organization consists of a network of reactions, finely regulated by their highly complex material components (enzymes), and regenerated by the very network that they control in an organizationally closed way. Metabolic organization is not globally distributed and functionally undifferentiated, in contrast to dissipative structures and many other examples of physical and chemical networks. Instead, living systems possess and generate hierarchical levels, differentiated regulatory mechanisms, and temporally decoupled processes, each of them contributing differently to the global maintenance of the system (Wimsatt [2002]).

Consider, for instance, a biological cell. In a cell, different localizable structures, generated within the system, make differentiable contributions to self-maintenance. The membrane—as a whole, integrated structure made out of many different molecules—acts as a constraint that (among other things) ensures adequate internal concentrations of metabolites. In turn, DNA (among other things) acts as a template for the synthesis of the proteins, which have to be continuously renewed due to their high rate of decay. In a word, the cell possesses different *parts*, produced within and by the system, that contribute differently to the maintenance of the system itself. In this way, functional attributions to each part are grounded.

The understanding of metabolic networks as organizationally closed and differentiated systems was put forward several decades ago by pioneers such as Rosen ([1973], [1991]), Maturana and Varela ([1973]), Varela *et al.* ([1974]), Varela ([1979]), Pattee ([1973], [1982]), and Gánti ([1975], [2003]). Beyond the specific (and sometimes simplistic) formulations of these authors, this approach is now increasingly accepted and has been developed in various scientific fields, including theoretical biology, biochemistry, and synthetic biology (some recent examples are Luisi [1998], [2006]; Kauffman [2000]; Cornish-Bowden *et al.* [2007]; Hofmeyr [2007]).

Insofar as their metabolic organization can be described as a closed and differentiated self-maintaining network, living systems possess all the relevant

complex) kind of catalyst reproducing copies of itself. In this case, functional attributions are not meaningful, since, here, the set makes a unique contribution to its maintenance, i.e., it produces copies of the same catalyst. Let us now suppose that some catalysts, instead of producing other copies, start producing molecules of a new kind, which generate, say, a new structure enclosing the network (a membrane). Suppose also that this new structure were to exert a constraint, by contributing, for instance, to ensuring a suitable concentration of raw materials within the system. We argue that, since this new system would generate different components (catalysts and membrane molecules) making different contributions to the maintenance of the system, it would be possible to attribute functions to them.



properties underlying functional analysis from the organizational perspective we are proposing here.<sup>9</sup>

## 5 Functions

Let us now focus on functions, by providing an explicit definition in organizational terms. According to the OA, a trait type *T* has a function if, and only if, it is subject to organizational closure *C* in a differentiated self-maintaining system *S*.

This definition implies the fulfilment of three different conditions. Accordingly, a trait *T* has a function if and only if:

- $C_1$ : *T* contributes to the maintenance of the organization *O* of *S*;
- $C_2$ : *T* is produced and maintained under some constraints exerted by *O*;
- $C_3$ : *S* is organizationally differentiated.

Accordingly, the heart has the function of pumping blood since pumping blood contributes to the maintenance of the organism by allowing blood to circulate, which in turn enables the transport of nutrients to and waste away from cells, the stabilization of body temperature and pH, and so on. At the same time, the heart is produced and maintained by the organism, whose overall integrity is required for the ongoing existence of the heart itself. Lastly, the organism is organizationally differentiated, since it produces numerous other structures contributing in different ways to the maintenance of the system.

It should be emphasized that the characterization of functions we are proposing is consistent with the one proposed by Wright. In this example, the heart is there because it pumps blood (otherwise the organism, and thus the heart, would disappear), and pumping blood is a consequence of the heart's being there. This consistency stems from the fact that the organizational account, by appealing to a causal loop at work in the organization of the system, provides an argument for naturalizing both the teleology and normativity of functions, which, at an organizational level, mirrors the explanatory strategy adopted by the etiological approaches. The resulting account represents an integration of the etiological and dispositional perspectives, since it may at the same time explain the existence of the trait and its current contribution to the self-maintenance of the system.

It should be noted that, in principle, functional attributions may concern either the constitutive *traits* of a differentiated self-maintaining system or the

<sup>9</sup> In this paper, we do not discuss the question of whether the class of organizationally closed and differentiated systems coincides with the class of living systems. Consequently, we leave open the possibility that, at least in principle, functional attributions may concern 'non-living' systems. In practice, however, all known cases of closed and differentiated self-maintaining systems *are* living systems.

*processes* that these structures concur to produce. In what follows, however, we will mainly refer to the functions of traits, which are consistent with the typical use of functional attributions in the relevant literature, as well as in ordinary language (see also Wimsatt [2002], p. 179).

In the remainder of this section, we will discuss in more detail some relevant implications of  $C_1$  and  $C_2$  above.

### 5.1 $C_1$ : Contributing to the maintenance of the organization

With respect to the contribution of T to O,  $C_1$  has two possible variants:

- (1) T contributes to the maintenance of O in the sense that, without T, S would cease to exist (i.e., T makes an indispensable contribution to O);
- (2) T contributes to the maintenance of O in the sense that *that specific organization* would not exist without T, but S could still maintain itself by modifying its organization (i.e., T does not make an indispensable contribution to O).

The existence of these two variants clarifies the distinction between ‘being subject to closure’ and ‘being indispensable’ in a given self-maintaining system. In Section 4.1 above we argued that organizational closure grounds normativity in the sense that the constitutive processes of a self-maintaining system ‘must’ work in a specific way, otherwise the organization itself would cease to exist. In fact, this general claim has to be further refined.

Self-maintaining systems can be grouped into classes (and possibly sub-classes) according to the kind of organization they possess. In this way, we can distinguish between flames and hurricanes, between animals and plants, between humans and cats, and so on. Whereas each class is identified by a minimal set of common distinctive processes and constraints that the system has to instantiate, individual members may possess a more complex organization, including more processes, constraints, and capacities. We call *regime of self-maintenance* each possible specific organization that an individual member of a class can adopt without ceasing to exist or losing its membership of that class. Each class may thus include several regimes of self-maintenance. In organizational terms, if a trait is subject to closure (and thus has a function), then *the specific regime of self-maintenance* that the system has adopted requires the said trait as an indispensable component. Nevertheless, not every functional trait contributes to all possible regimes of self-maintenance of a given class, which means that an individual system can sometimes compensate for the breakdown of a component by shifting to a different regime of self-maintenance, in which the defective trait is no longer required. In contrast, some functional traits are

indispensable, in that they are required for all regimes of self-maintenance that a member of a class could possibly adopt.

To exemplify this distinction, one may consider two typical biological functions in humans: (i) the heart's pumping of blood and (ii) the eyes' transduction of light. In the first case, the functional trait is indispensable, because it contributes to generating a global process (the circulation of blood), which is required in order to preserve the existence of this class of systems, whatever regime of self-maintenance of the members is considered. In this sense, *there are no organizational alternatives* to blood pumping for humans to be viable. In the second case, in contrast, the transduction of light contributes to generating a capacity (seeing), which constrains other processes in specific modes of self-maintenance but is not indispensable for human beings (blind people can survive). However, since the transduction of light is functional, this crucially means that a whole network of processes depends in some way on the capacity of the eyes to transduce light. Accordingly, if the eyes were to stop performing their function or if they were to malfunction, a global constraint (vision, in this case) would disappear and the system would be forced to shift to a new regime of self-maintenance (in this case, find new ways of finding food, moving around; etc.).<sup>10</sup> To the extent that, as we have already pointed out (see footnote 5), self-maintenance is a context-dependent determination, the same caveat applies of course to the distinction between 'simply functional' and 'indispensable' traits. Nevertheless, we maintain that, for each given individual in a given environment, the distinction usefully clarifies a fundamental aspect about the way in which a trait can be said to contribute to the self-maintenance of a system.

## 5.2 C<sub>2</sub>: Producing the functional trait

With respect to the production and maintenance<sup>11</sup> of T by O, an important distinction stems from the specific regime of self-maintenance producing the functional trait.

As already discussed, since not all functional processes are indispensable, a system may adopt more or less complex regimes of self-maintenance, following

<sup>10</sup> Of course, the case in which the breakdown of a trait implies the complete loss of a given capacity is an extreme case. In many cases, the breakdown of a functional trait simply requires a shift to a new regime of self-maintenance in which the system will be able to perform the same set of functions through a (more or less) different physical organization. This is typically the case with the brain and neural networks, in which physically different parts can subserve the same capacity.

<sup>11</sup> The specific production relation at work between the whole organization and the entity being produced could provide a relevant criterion to distinguish between biological traits and artefacts. In particular, biological traits would correspond to traits being at the same time produced *and* maintained by the ongoing organization of the system. Artefacts, in contrast, would correspond to entities that, although produced by a self-maintaining organization, do not depend on it for their ongoing maintenance.

the complexity of the ongoing organization. Accordingly, we submit the following definitions. For each given class of self-maintaining systems, the *primary* function  $F_p$  of  $T$  is the contribution of  $T$  to the self-maintenance of  $S$  that is subject to closure in the more basic regime of self-maintenance. In contrast, a *secondary* function  $F_s$  of  $T$  is the contribution of  $T$  to the self-maintenance of  $S$  that is subject to closure in whatever (comparatively) more complex regime of self-maintenance.

For instance, the heart has the primary function of pumping blood because pumping blood is the contribution of the heart subject to closure in the more basic regime of self-maintenance required for its own production and maintenance. In contrast, the heart might have the secondary function to, say, make a characteristic noise, since this effect could possibly contribute (by permitting medical diagnoses, for instance) to a more complex regime of self-maintenance (which in this case would include socio-technical interactions), which would also help to produce and maintain the heart. Similarly, the nose has the primary function to warm and humidify air, whereas it may have the secondary function to support eyeglasses, since the complexity of the regime of self-maintenance in which the first contribution is subject to closure is lower than that of the second one.

In most cases, in which no distinct regimes of self-maintenance are available in a given class, primary and secondary functions simply collapse into a unique definition. For those classes admitting distinct regimes of self-maintenance, the organizational approach explicitly admits the possibility of a given trait performing simultaneously more than one function. In such cases, however, a distinction between different 'levels' of functionality would be preserved and, in particular, the concept of primary function would correspond to the etiological notion of *proper* function (see Section 6.3 below).

At first sight, the primary function of a trait can be seen as being more 'fundamental', in the sense that, usually, secondary functions are performed in addition to a trait's primary function. Or, more precisely, the organization in which secondary functions are subject to closure usually includes the organization of the primary function. For instance, the organization in which the nose performs the function of supporting eyeglasses includes the organization in which it has the proper function of warming and humidifying air. It should be noted, however, that the notions of primary and indispensable functions are conceptually *independent*, and that their mutual relations and dependencies have to be clarified case by case. A trait's primary function may be indispensable in a given class (as in the case of the heart) or not (as for the eyes). Therefore, a trait may not only possess simultaneously a primary function and one or more secondary functions, it could also perform a secondary function (i.e., supporting eyeglasses) without actually performing or being able to perform the primary function (i.e., warming and humidifying air). Whatever the specific

relations between the multiple functions of a trait might be, the central point is that all functional attributions to a trait  $T$ , be they primary or secondary, provide an answer to both the question ‘why  $T$ ?’ and the question ‘what is  $T$  for?’ in different regimes of self-maintenance.

Again, it is worth clarifying that the distinction between primary and secondary function is a context-dependent one. What counts as a primary or secondary function of a trait depends on what counts as a basic regime of self-maintenance, which in turn depends on the context with respect to which the class of self-maintaining systems (and its admissible regimes) has been defined. If the context changes, the definition of the constitutive organization of the class may possibly change. Nevertheless, we hold that, for each specifiable environment or context, the distinction between primary and secondary is not arbitrary and captures a relevant property of complex biological systems, in which some categories of structures (such as proteins or cerebral areas) may perform several functions, and regularly shift between them.

## 6 Implications and Objections

Before concluding, we will discuss some relevant implications of the OA, as well as some objections that could possibly arise.

### 6.1 Functional *versus* useful

A first relevant implication of the OA is that it possesses the conceptual resources to distinguish not only between two levels of functionality, but also between functionality and usefulness. Whereas functions are the class of contributions satisfying the three conditions formulated in Section 5, we argue that what are usually called ‘useful’ contributions are those produced by entities that satisfy  $C_1$  and  $C_3$  but not  $C_2$ . An entity is useful if it contributes to the maintenance of an organizationally closed and differentiated system, without being produced under some constraint generated by the system. Accordingly, various entities such as oxygen, food, and gravity are useful for a given class of self-maintaining systems, without being functional. In the same way as functional entities, useful entities may be indispensable or not, depending on the specific regime of self-maintenance to which they contribute. In addition, since they satisfy  $C_1$  but not  $C_2$ , useful entities are subject to the same kind of normativity as functional entities (one can answer the question ‘What is the oxygen for in an organism?’), whereas they do not refer to a teleological dimension (the relation of useful entities with the self-maintaining organization does not allow one to answer the question ‘Why do they exist in that system?’).

## 6.2 Dysfunctions, side effects, and accidental contributions

A second relevant implication of the OA is that it provides a principled criterion to distinguish between functional and dysfunctional processes and to rule out functional attributions to irrelevant effects.

In each class of self-maintenance, a trait may perform one primary function and possibly several secondary functions according to the specific regime of self-maintenance in which it is subject to closure. Dysfunctions appear whenever a trait fails to adequately perform its primary and/or secondary function. A dysfunctional trait is a trait that fits  $C_2$  and  $C_3$ , but fails to fit  $C_1$ . When a dysfunction occurs, the system has to react in order either to compensate for the dysfunction and maintain the current regime of self-maintenance, or to shift to a new one if the function is not indispensable. Dysfunctions are defined in relation to a specific regime of self-maintenance in a specific class of self-maintaining systems. Hence, to interpret some neutral or deleterious effect of a trait in an individual system as dysfunctional, the current organization of the individual has to belong to a specifiable regime of self-maintenance in a class of self-maintaining systems in which the function of the trait is defined.

In turn, irrelevant effects of traits are effects that, while they satisfy  $C_2$  and  $C_3$ , do not contribute to maintaining the organization of the system (and consequently do not fit  $C_1$ ) without being dysfunctional, in the sense that the irrelevant effect *does not prevent* the trait from performing its function(s).

Finally, the OA allows one to interpret the concept of 'accidental' contributions. A contribution to self-maintenance is accidental if it occurs only in relatively infrequent and occasional circumstances. In the OA, however, the frequency of a contribution to self-maintenance is irrelevant for determining its status. Thus, both functional and useful contributions might be systematic or accidental, following the specific context with respect to which the relevant class of self-maintaining systems has been defined.

## 6.3 Proper functions and selected effects

An interesting issue raised by the OA is the relation between its concept of primary functions and the etiological concept of proper functions. The problem is quite complex, and we leave a detailed analysis of it for an independent paper. However, the heart of the issue is as follows: we suggest that the attributions of primary functions in organizational terms will in many cases coincide with attributions of proper functions in selective terms, even if diverging cases exist. As we have already discussed, primary functions in organizational terms are the processes produced by a trait  $T$  that are subject to closure in the (comparatively) more basic organization required to produce and maintain  $T$ . In this sense, it is likely that the basic organization in which a trait  $T$  is subject to closure is in

fact the oldest one in evolutionary terms, if compared to more complex organizations involving a higher number of constraints and constitutive processes. For instance, to take an obvious example, the organization in which the process ‘transducing light’ is subject to closure is evolutionarily older than that in which the process ‘supporting eyeglasses’ is subject to closure. Since more basic organizations are presumably the oldest in evolutionary terms, traits subject to closure in those organizations are more likely to have been subject to a selective process, and to have made functional contributions to ‘selected effects’.

It should be emphasized that convergence between these two characterizations does not imply their theoretical equivalence. Primary functions, we maintain, are such on the basis of the normative and teleological dimensions generated by the closed organization of a differentiated self-maintaining system. The fact that some of these functions are also selective effects is, as such, irrelevant to an organizational interpretation of functional attributions.

Despite the convergence, attributions of primary or proper functions from an organizational or etiological perspective may differ in at least two main cases. The first case concerns those traits that have been selected for their effects, but that no longer perform their proper function. Here, the SE theories would still attribute a proper function to the trait, whereas the organizational account would not, and would instead claim that the now-useless trait *had* a proper function in the past and now does not. The second case is the symmetrical opposite, in which a trait contributes to producing a process subject to closure in the more basic self-maintaining organization required to produce T. Here, the OA would attribute function to the trait, whereas the SE theories would not. It is on the basis of this second divergence that, as we emphasized in Section 5, the organizational approach may provide a solution for the epiphenomenalism of the SE theories, while still accounting for the teleological dimension of functions.

## 6.4 Reproduction

A possible objection to the OA is that it seems to fail to account adequately for the functions of reproductive traits. Reproductive traits, the objection would charge, have a function in biological systems, even if they do not appear to be subject to closure since they do not contribute to the self-maintenance of the system producing and maintaining them. Again, the issue would require an independent paper, developing an articulated theoretical analysis. Nevertheless, let us develop here the gist of the argument.

The objection is right in pointing out that reproductive traits are not subject to closure in *any individual* self-maintaining system. Indeed, reproductive traits are neither subject to closure in the reproducing system nor in the reproduced system, which do not produce and maintain them. In this sense, we explicitly

claim that, in the organizational account, reproductive traits are not functional with respect to the individual self-maintaining system.

Yet, the objection ignores an alternative explanation. The OA may explain functional attributions to reproductive traits by appealing to ‘second order’ self-maintaining systems, composed of a set of individual self-maintaining systems, in which reproductive traits could be subject to organizational closure. As a matter of fact, organisms belong to higher-order *historical* systems, in which individuals, each having a limited lifetime, are continuously reproduced through generations. Similarly, organisms are components of higher-level *ecological* systems, i.e., networks of interactions among organisms that constrain some of the conditions of existence of each individual. In both cases, a form of organizational closure seems to exist between the individual organisms and the higher-order system, be it historical, ecological, or both. Individuals contribute to the maintenance of the higher-order system, which in turn reproduces and maintains them. In this sense, both kinds of higher-level systems are candidates as second-order self-maintaining systems, of which individuals would constitute functional parts. In particular, reproductive traits of individuals would be subject to closure in the organization of the second-order systems, and then have a function.

Of course, an articulated theory of second-order self-maintaining systems is still to be formulated. In this sense, the argument we have developed here still lacks adequate justification. Yet, we maintain that the OA has the theoretical resources to correctly capture the *sui generis* nature of reproductive functions in the whole set of biological functions. An adequate explanation of functional ascriptions to reproductive traits constitutes, then, one of the main objectives of the OA’s research programme.

## 6.5 Relation with other ‘unitarian’ approaches

In the recent philosophical literature about functions, a number of proposals have tried to unify the dispositional and etiological approaches into a common framework. For instance, Griffiths ([1993]) suggests that the etiological approach can be incorporated into Cummins’ systemic approach. In particular, ‘the proper functions of a biological trait are the functions it is ascribed in a functional analysis of the capacity to survive and reproduce (fitness) which has been displayed by animals with that feature’, which amounts to affirming that ‘the proper functions of a trait are those effects of the trait that were components of the fitness of ancestors’ (Griffiths [1993], pp. 146–7). Kitcher ([1993]), Walsh ([1996]), and Walsh and Ariew ([1996]) have put forward approaches based on similar strategies. Here, we focus on Walsh’s proposal, leaving a more articulated discussion of the relation between these ‘non-historical’ approaches (Wouters [2005]) and the organizational account for an independent paper.



Walsh proposes a relational theory of functions, according to which ‘the/an evolutionary function of a token of type X with respect to selective regime R is to m if and only if X’s doing m positively (and significantly) contributes to the average fitness of individuals possessing X in R’ (Walsh and Ariew [1996], p. 263). According to whether the relevant selective regime is past or current, the function will be past or current.

Indeed, the relational approach succeeds in providing a non-historical definition of functions, which can apply to both historical and current functions. However, the conceptual strategy adopted is the opposite of the one we propose in this paper. As McLaughlin points out ([2001], pp. 128–31), the relational theory unifies historical and current functions by interpreting past functions as past contributions to fitness. The etiological approach is then reinterpreted as a propensity view applied to past selective regimes. The main consequence is that functional attributions to a trait can no longer explain its current presence, which is the main explanandum of etiological theories. The relational theory may either appeal to the function of past traits to explain the current presence of a trait, or to the current function of a trait to explain the existence of future traits. But functional attributions are no longer able to account for the existence of the current function bearer.

The organizational approach adopts the opposite strategy in that it aims to provide a unified definition of functions by extending the teleological dimension to the current activity of a trait. Functional attributions to both past and current traits explain the presence of the trait in terms of the effects of its contribution to the self-maintenance of the system to which it belongs. The organizational concept of function applies to classes of self-maintaining systems in current or past regimes of self-maintenance, by preserving in both cases its teleological and normative dimensions.

## 7 Conclusions

In this paper, we have put forward an organizational account in which biological functions are defined as causal relations subject to closure in living systems, interpreted as the most typical examples of organizationally closed and differentiated self-maintaining systems. We have argued that this account justifies the grounding of the teleological and normative dimensions of functions in the current organization of the system, in that it provides an explanation for the existence of the function bearer and, at the same time, identifies in a non-arbitrary way the norms that functions are supposed to obey. Accordingly, we have suggested that the OA may combine the etiological and dispositional perspectives in an integrated theoretical framework.

Some authors have recently put forward accounts of biological functions closely related to the approach developed in this paper. In particular,

Christensen and Bickhard ([2002]) have suggested, relying on their own work on the notion of biological autonomy, that the organization of autonomous systems provides an adequate grounding for the normativity of functional attributions. In a similar vein, McLaughlin ([2001]) has developed an account in which both the teleology and normativity of functions can be naturalized in the organization of self-reproducing systems. Despite some terminological differences, the central idea of these approaches (i.e., that the organizational closure instantiated by living systems provides an adequate basis to naturalize functions) fundamentally coincides with that defended in this paper, and we explicitly recognize the theoretical relationship.

Nevertheless, we believe that these authors did not succeed in building a fully satisfactory account of this central idea, which could (at least in part) explain why their work has scarcely been echoed in the philosophical literature about functions. Our account may represent a substantial philosophical improvement on previous similar approaches in two aspects. First, it articulates a fine-grained description of its central concepts and an explicit definition of biological functions, thereby placing precise constraints on functional attributions. Second, it spells out—at least in a preliminary way—the major implications of the central idea with respect to several issues with which a theory of function is supposed to deal.

Of course, most of the philosophical work is still to be done. Nevertheless, the arguments developed in this study may constitute a useful step towards a unifying organizational theory of functions for living systems and artefacts.

### **Acknowledgements**

The authors wish to thank Jean Gayon, Philippe Huneman, and Xabier Barandiaran for valuable feedback on earlier versions of this paper. The work was funded by Centre National de Recherche Scientifique, France (ACI programme 'La notion de fonction dans les sciences humaines, biologiques et médicales' to MM); Ministerio de Ciencia e Innovación, Spain ('Juan de la Cierva' program to MM); Ministerio de Ciencia e Innovación, grant no. HUM2005-02449 and BFU2006-01951/BMC; Universidad del País Vasco (postdoctoral fellowship to MM); Gobierno Vasco (predoctoral fellowship to CS and grant no. 937 IT-250-07); and the European Union (Feder funds).

Matteo Mossio

*Institut d'Histoire et de Philosophie des Sciences et des Techniques  
CNRS/ Université Paris1/ ENS, 13 rue du Four, 75006*

*Paris, France*

*Matteo.Mossio@ens.fr*

Matteo Mossio, Cristian Saborido, Alvaro Moreno  
*Department of Logic and Philosophy of Science*  
*University of the Basque Country*  
*Avenida de Tolosa 70, 20080*  
*San Sebastian, Spain*  
*cristian.saborido@ehu.es*  
*alvaro.moreno@ehu.es*

## References

- Adams, F. R. [1979]: 'A Goal-State Theory of Function Attributions', *Canadian Journal of Philosophy*, **9**, pp. 493–518.
- Allen, C. and Bekoff, M. [1995]: 'Biological Function, Adaptation and Natural Design', *Philosophy of Science*, **62**, pp. 609–22.
- Allen, C., Bekoff, M. and Lauder, G. V. (eds) [1998]: *Nature's Purposes*, Cambridge, MA: MIT Press.
- Ariew, A. R., Cummins, R. and Perlman, M. (eds) [2002]: *Functions*, Oxford: Oxford University Press.
- Barandiaran, X. and Moreno, A. [2008]: 'Adaptivity: From Metabolism to Behavior', *Adaptive Behavior*, **16**, pp. 325–44.
- Bedau, M. A. [1992]: 'Goal-Directed Systems and the Good', *The Monist*, **75**, pp. 34–49.
- Bickhard, M. H. [1993]: 'Representational Content in Humans and Machines', *Journal of Experimental and Theoretical Artificial Intelligence*, **5**, pp. 285–333.
- Bickhard, M. H. [2000]: 'Autonomy, Function, and Representation', *Communication and Cognition – Artificial Intelligence*, **17**, 3–4, pp. 111–31.
- Bigelow, J. and Pargetter, R. [1987]: 'Functions', *Journal of Philosophy*, **84**, pp. 181–96. Reprinted in Buller [1999], pp. 97–114.
- Boorse, C. [1976]: 'Wright on Functions', *Philosophical Review*, **85**, pp. 70–86.
- Boorse, C. [2002]: 'A Rebuttal on Functions', in Ariew et al. [2002], pp. 63–112.
- Buller, D. J. (ed.) [1999]: *Function, Selection, and Design*, Albany, NY: SUNY Press.
- Canfield, J. [1964]: 'Teleological Explanation in Biology', *British Journal for the Philosophy of Science*, **14**, pp. 285–95.
- Chandrasekhar, S. [1961]: *Hydrodynamic and Hydromagnetic Stability*, Oxford: Clarendon.
- Christensen, W. D. and Bickhard, M. H. [2002]: 'The Process Dynamics of Normative Function', *The Monist*, **85**, 1, pp. 3–28.
- Cornish-Bowden, A., Cárdenas, M. L., Letelier, J. C. and Soto-Andrade, J. [2007]: 'Beyond Reductionism: Metabolic Circularity as a Guiding Vision for a Real Biology of Systems', *Proteomics*, **7**, pp. 839–45.
- Craver, C. F. [2001]: 'Role Functions, Mechanisms, and Hierarchy', *Philosophy of Science*, **68**, pp. 53–74.
- Cummins, R. [1975]: 'Functional Analysis', *Journal of Philosophy*, **72**, pp. 741–65. Reprinted in Buller [1999], pp. 57–83.
- Cummins, R. [2002]: 'Neo-Teleology', in Ariew et al. [2002], pp. 157–72.

- Davies, P. S. [1994]: 'Troubles for Direct Proper Functions', *Noûs*, **28**, pp. 363–81.
- Davies, P. S. [2000]: 'Malfunctions', *Biology and Philosophy*, **15**, pp. 19–38.
- Davies, P. S. [2001]: *Norms of Nature. Naturalism and the Nature of Functions*, Cambridge, MA: MIT Press.
- Field, R. J., Körös, E. and Noyes, R. M. [1972]: 'Oscillations in Chemical Systems. II. Thorough Analysis of Temporal Oscillation in the Bromate-Cerium-Malonic Acid System', *Journal of the American Chemical Society*, **94**, pp. 8649–64.
- Field, R. J. and Noyes, R. M. [1974]: 'Oscillations in Chemical Systems. IV. Limit Cycle Behavior in a Model of a Real Chemical Reaction', *Journal of Chemical Physics*, **60**, pp. 1877–84.
- Gánti, T. [1975]: 'Organisation of Chemical Reactions into Dividing and Metabolizing Units: The Chemotons', *BioSystems*, **7**, pp. 15–21.
- Gánti, T. [2003]: *The Principles of Life*, With a commentary by J. Griesemer and E. Szathmáry (eds), Oxford: Oxford University Press.
- Glansdorff, P. and Prigogine, I. [1971]: *Thermodynamics of Structure, Stability and Fluctuations*, London: Wiley.
- Godfrey-Smith, P. [1994]: 'A Modern History Theory of Functions', *Noûs*, **28**, pp. 344–62. Reprinted in Buller [1999], pp. 199–220.
- Griffiths, P. E. [1993]: 'Functional Analysis and Proper Functions', *British Journal for the Philosophy of Science*, **44**, pp. 409–422. Reprinted in Buller [1999], pp. 143–58.
- Hardcastle, V. G. [2002]: 'On the Normativity of Functions', in Ariew *et al.* [2002], pp. 144–56.
- Hempel, C. G. [1959]: 'The Logic of Functional Analysis', in L. Gross (ed.), *Symposium on Sociological Theory*, New York: Harper and Row, pp. 271–87.
- Hofmeyr, J.-H. S. [2007]: 'The Biochemical Factory That Autonomously Fabricates Itself: A Systems Biological View of the Living Cell', in F. Boogerd, F. J. Bruggeman, J.-H. S. Hofmeyr and H. V. Westerhoff (eds), *Systems Biology: Philosophical Foundations*, Amsterdam: Elsevier, pp. 217–42.
- Kauffman, S. [2000]: *Investigations*, Oxford: Oxford University Press.
- Kitcher, P. [1993]: 'Function and Design', *Midwest Studies in Philosophy*, **18**, pp. 379–97.
- Luisi, P. L. [1998]: 'About Various Definitions of Life', *Origins of Life and Evolution of Biospheres*, **28**, pp. 613–22.
- Luisi, P. L. [2006]: *The Emergence of Life: From Chemical Origins to Synthetic Biology*, Cambridge: Cambridge University Press.
- Maturana, H. and Varela, F. J. [1973]: *De máquinas y seres Vivos—Una teoría sobre la organización biológica*, Santiago, Chile: Editorial Universitaria S.A.
- McLaughlin, P. [2001]: *What Functions Explain. Functional Explanation and Self-reproducing Systems*, Cambridge: Cambridge University Press.
- Melandier, P. [1997]: *Analyzing Functions. An Essay on a Fundamental Notion in Biology*, Stockholm: Almqvist & Wiksell International.
- Millikan, R. G. [1984]: *Language, Thought, and Other Biological Categories*, Cambridge, MA: MIT Press.
- Millikan, R. G. [1989]: 'In Defense of Proper Functions', *Philosophy of Science*, **56**, pp. 288–302.

- Millikan, R. G. [1993]: 'Propensities, Exaptations, and the Brain', in R. G. Millikan (ed.), *White Queen Psychology and Other Essays for Alice*, Cambridge, MA: MIT Press, pp. 31–50.
- Millikan, R. G. [2002]: 'Biofunctions: Two Paradigms', in Ariew et al. [2002], pp. 113–43.
- Nagel, E. [1961]: *The Structure of Science*, London: Routledge & Kegan Paul.
- Nagel, E. [1977]: 'Teleology Revisited', *Journal of Philosophy*, **74**, pp. 261–301.
- Neander, K. [unpublished]: 'Teleology in Biology', Paper presented to AAP conference 1980.
- Neander, K. [1991]: 'Function as Selected Effects: The Conceptual Analyst's Defense', *Philosophy of Science*, **58**, pp. 168–84.
- Nicolis, G. [1989]: 'Physics of Far-from-Equilibrium Systems and Self-organisation', in P. Davies (ed.), *The New Physics*, Cambridge: Cambridge University Press, pp. 316–47.
- Nicolis, G. and Prigogine, I. [1977]: *Self-organisation in Non-equilibrium Systems: From Dissipative Structures to Order through Fluctuation*, New York: Wiley.
- Nissen, L. [1980]: 'Nagel's Self-regulation Analysis of Teleology', *Philosophical Forum*, **12**, pp. 128–38.
- Pattee, H. H. [1973]: 'The Physical Basis and Origin of Hierarchical Control', in H. H. Pattee (ed.), *Hierarchy Theory*, New York: Braziller, pp. 73–108.
- Pattee, H. H. [1982]: 'Cell Psychology: An Evolutionary Approach to the Symbol-Matter Problem', *Cognition and Brain Theory*, **4**, pp. 325–41.
- Rosen, R. [1973]: 'On the Dynamical Realizations of (M, R)-Systems', *Bulletin of Mathematical Biophysics*, **35**, pp. 1–9.
- Rosen, R. [1991]: *Life Itself: A Comprehensive Inquiry into the Nature, Origin and Fabrication of Life*, New York: Columbia University Press.
- Rosenblueth, A., Wiener, N. and Bigelow, J. [1943]: 'Behavior, Purpose and Teleology', *Philosophy of Science*, **10**, pp. 18–24.
- Ruiz-Mirazo, K. [2001]: *Physical Conditions for the Appearance of Autonomous Systems with Open-ended Evolutionary Capacities*, Ph.D. dissertation, University of the Basque Country.
- Ruse, M. [1971]: 'Functional Statements in Biology', *Philosophy of Science*, **38**, pp. 87–95.
- Ruse, M. [2003]: *Darwin and Design: Does Evolution have a Purpose?*, Cambridge, MA: Harvard University Press.
- Sommerhoff, G. [1950]: *Analytical Biology*, Oxford: Oxford University Press.
- Sommerhoff, G. [1959]: 'The Abstract Characteristics of Living Organisms', in F. E. Emery (ed.), *Systems Thinking*, London: Harmondsworth, pp. 147–202.
- Varela, F. J. [1979]: *Principles of Biological Autonomy*, New York: North Holland.
- Varela, F. J., Maturana, H. and Uribe, R. [1974]: 'Autopoiesis: The Organisation of Living Systems, Its Characterization and a Model', *BioSystems*, **5**, pp. 187–96.
- Walsh, D. M. [1996]: 'Fitness and Function', *British Journal for the Philosophy of Science*, **47**, pp. 553–74.

- Walsh, D. M. and Ariew, A. R. [1996]: 'A Taxonomy of Functions', *Canadian Journal of Philosophy*, **26**, pp. 493–514. Reprinted in Buller [1999], pp. 257–79.
- Wimsatt, W. [2002]: 'Functional Organisation, Functional Inference, and Functional Analogy?', in Ariew *et al.* [2002], pp. 174–221.
- Wright, L. [1973]: 'Functions', *Philosophical Review*, **82**, pp. 139–68.
- Wright, L. [1976]: *Teleological Explanations: An Etiological Analysis of Goals and Functions*, Berkeley: University of California Press.
- Wouters, A. G. [2005]: 'The Function Debate in Philosophy', *Acta Biotheoretica*, **53**, pp. 123–51.