Agential Teleosemantics

A PhD Thesis by:

Tiago Rama

Supervisor:
Sergio Balari

PhD program in Cognitive Science & Language
Facultat de Filosofia i Lletres
Universitat Autònoma de Barcelona

October 2022

Abstract

The field of the philosophy of biology is flourishing in its aim to evaluate and rethink the view inherited from the previous century —the Modern Synthesis. Different research areas and theories have come to the fore in the last decades in order to account for different biological phenomena that, in the first instance, fall beyond the explanatory scope of the Modern Synthesis. This thesis is anchored and motivated by this revolt in the philosophy of biology.

The central target in this context is the possibility of naturalizing teleology, a classical nightmare for the history of biology itself. This requires, principally, understanding the causes of teleological explanations without assuming an unfashioned backward causation of sorts. As the riddles of teleological explanations are about their temporal dimension, I analyze different temporal scales of biological processes: evolutionary, developmental, and physiological.

The first one is the one defended in the context of the Modern Synthesis. As expected, one of the aims of this thesis is to evaluate the adequacy of an evolutionary account of teleological explanation. The scrutiny is negative. Evolutionary explanations in the context of the Modern Synthesis lack the necessary causal roots to naturalize teleology. Concerning the physiological scale, a long tradition pushed up by Kant and the organicist movement in the 20^{th} century allows us to better understand how teleological explanations can be naturalized in physiological process. The key notions in this temporal scale are self-organization and the recursive, looped character of physiological process. While the physiological scale may be suitably accounted by contemporaries views, such as Autonomous Systems Theory, different central teleological phenomena remain unexplained from a purely physiological perspective. In particular, different issues concerning the (adaptive) construction of organism —such as plasticity, robustness, variation, novelty, inheritance— deserve an ontogenetic analysis.

The principal aim of this thesis is to provide a theory of teleological development that falls beyond the Modern Synthesis' framework and is prompted by different insights from the history of biology. I call it *Agential Teleosemantics*. It rests on two central pillars. First, that developmental processes, beyond any gene-centered stance, can be understood in informational terms; i.e. developmen-

tal processes are about the interaction of developmental resources conveying *biological information*. The second ingredient is agentivity, namely the idea that development is regulated by an *agentive system* according to the adaptiveness of the phenotypic outcomes produced. The role of agency in Agential Teleosemantic is equivalent to the role of genes in the Modern Synthesis: it is responsible for explaining the order and the adaptive complexity in the living realm.

The second target of this thesis regards the possibility of naturalizing intentional explanations in cognitive science. The central project involved in such an aim is known as teleosemantics. Classical teleosemantics however is etiological: it explains the teleofunctions of representational systems in terms of evolutionary processes. The different disputes in the contemporary philosophy of biology provide two insights to analyze teleosemantics in cognitive science. First, the challenges against the Modern Synthesis must be extended to the evolutionary approach of etiological teleosemantics. Second, as Agential Teleosemantics suggests an alternative source for teleofunctions —ontogeny, I offer an attempt to integrate Agential Teleosemantics into cognitive science in order to provide an alternative teleosemantic project to understand intentional explanations in cognitive science.

Declaration

I declare that this thesis is my own original work. Up to the best of my knowledge, all the works done by others used in this thesis are correspondingly referred.

Tiago Rama Folco

Acknowledgments

I thank Sergio Balari for his support throughout these years, and all the conversations about philosophy, biology, and many other topics. Many of the ideas presented here were the result of his guiding comments, suggestions, and many hours of thinking together. I would have never come to know the exciting field of the philosophy of biology without meeting him. But most important, his was the perfect tutoring for my way of working and doing philosophy, and this made me enjoy my years as a PhD student. Fortunately, doing this thesis was not a torturous route and this is due to the way we worked together. I am aware that this is valuable and I am really grateful for that. Last, but not least, he helped me a lot in turning this thesis into something closer to readable English, so I thank him for carrying out this (hard) task.

I am also grateful with Xabier Barandiaran for receiving me at the IAS-Research Group, working together during several months, and for the many insightful comments and conversations that have contributed to many ideas presented in this thesis, particularly in chapter 6.

I thank the Centre de Lingüística Teòrica of the Universitat Autònoma de Barcelona for receiving me during these three years, providing me with a workspace, and for access to many bibliographical resources that have contributed to writing this thesis in optimal conditions.

This work has received support from the the National Agency of Investigation and Innovation (Uruguay) through grant POS-EXT-2018-1-154759, and from the Spanish Government through grant FFI2017-87699-P to the Centre de Lingüística Teòrica of the Universitat Autònoma de Barcelona.

To whom it may concern

It is true that Darwin, when considering natural selection, leaves out of account the *causes* which have produced the alterations in separate individuals, and deals in the first place with the way in which such individual deviations gradually become the characteristics of a race, variety or species. To Darwin it was of less immediate importance to discover these causes —which up to the present are in part absolutely unknown, and in part can only be stated in quite general terms— than to find a rational form in which their effects become fixed, acquire permanent significance. It is true that in doing this Darwin attributed to his discovery too wide a field of action, made it the sole agent in the alteration of species and neglected the causes of the repeated individual variations, concentrating rather on the form in which these variations become general; but this is a mistake which he shares with most other people who make any real advance.

Friedrich Engels 1878, 82-83

Note on the bibliography

In the bibliography, I treat the works of authors from the 18^{th}, the 19^{th}, and early 20^{th} centuries in the following way: in the case of those texts from which I use some quotes, I cite first the original edition and then refer to a more recent (and often translated) one. I use only the original edition for those texts that are referred to in the text. In the case of epigraphs, I refer to the original editions, and the pages are taken from recent editions, which are also added to the bibliography.

I thank Sergio Balari for translating into English some works by Karl Ernst von Baer. Some of the quotes used in this thesis are translated by him. For other authors whose works have not been translated into English, I refer to the texts from which I have taken the translated quotes.

Contents

List of Figures

List of Tables

Introduction

> Theoretical perspectives in science coordinate models and phenomena. Coordination is necessary because phenomena are complex, our scientific interests in them are heterogeneous, and the number of possible ways of representing them in models is large.
>
> James Griesemer 2000a, 348-349

> It is not sufficient to point out problems with a concept. It is as important to find an alternative that will be free of these problems and that will offer at least as fruitful a research program as the old perspective.
>
> Eva Jablonka 2004, 366

This thesis is about two salient philosophical issues in the Life Sciences. Here I try to offer a coherent yet not complete view of teleology and intentionality and I would like to start offering some introductory remarks on these notions.

Teleology

Teleological talk in biology is widespread in our everyday way of thinking about life and living beings. We say that a trait performs a particular function in the benefit of the organism possessing such a trait, such as hearts beating in order to pump blood throughout the body. We also say that the parts of the body of an organism and its behaviors have a certain purpose: hands have the purpose of catching things, bird's songs have the purpose of communicating different states of affairs to other birds, and so on. We also see in certain activities carried out by an organism attempts to fulfill a goal: tress growth in a certain way to capture more solar energy, bacteria move toward areas of their environments with a high concentration of nutrients in order to feed. Teleology is connected with all these notions: functions, purposes, and goals. Teleological explanations in biology explain a certain phenomenon by appealing to the functions, purposes, and goals involved in the phenomenon: tress grow in a particular way *because* this maximazes their intake of sunlight, birds sing to communicate with other birds *because*

communication is relevant for their survival, and hearts beat *because* doing so is essential to keep the system alive.

It is intriguing that teleological explanations appear to be paradigmatic of biology only. There appears to be no room for teleological locutions in such scientific fields as physics or chemistry and the use of functional or purposive talk to describe, for instance, the hydrologic cycle or the movement of planets is merely metaphoric or heuristic. Clouds have no functions; planets do not move for a purpose.

The first important issue to ask then is whether teleological talk in biology also has this heuristic or metaphorical flavor. Perhaps biology can dispense with teleological idioms, relegating them to the status of mere *façons de parler* that humans use to describe and interact with organisms in pre-theoretical contexts. Those philosophers who support this view are typically eliminativists. But eliminativism faces the difficulty of clearly identifying the differences that exist between living and non-living systems. If biology does not require any specific kind of explanation, how do we account for the distinct nature of living systems?

In this thesis, I will vindicate a non-eliminativist position, which straight away requires addressing a central (and classical) problem with teleology: the so-called Kant's Puzzle. Is it possible to find a legitimate epistemic place for teleological explanations in the natural sciences? This puzzle foregrounds a tension between what is explanatory useful (teleological idioms) and what is explanatory valid (according to the foundations of natural science). But, why is it the case that teleology is so problematic for the natural sciences? The main problem is that teleological explanations appear to invert the order of explanatory relations: future events (e.g. goals) explain current activity (means toward the goal), while the consensus view in modern science is that the world is a chain of events or processes in which one event is caused by an earlier one and causes future ones. Future events cannot cause past events. This inverts the structure of causal explanations and would force us to accept what is typically known as *backward causation*. So, how can teleological explanation be accommodated in the causal picture of the world accepted by modern science? Therefore Kant's Puzzle on teleology is a problem about causality, and explaining the causal structure of teleological explanations is one of the main aims of this thesis.

Intentionality

Intentionality is also a technical notion in science, but we constantly resort to to intentional explanations in our daily life. For instance, we say that someone goes to the supermarket because she *believes* that there she would find the food she needs. Someone goes on holiday because she *wishes* to take a break. People *think* about which would be the best restaurant to go to tonight and then make the

decision. Intentionality is related to these locutions: beliefs, desires, and thoughts. In general, intentionality pops up in any explanation that involves some sort of cognitive or psychological state.

Intentionality lies at the core of cognitive science. The main tenet of cognitive science is that intentional explanations that involve psychological states can be analyzed in scientific terms. Particularly, cognitive science was born with the idea that psychological states are constituted by central units of information: representations. Thus, one person represents the supermarket and the food she needs and then goes out for shopping; another person represents herself on a beach and then she books a holiday to the Greek Islands. Cognitive science is interested in explaining how cognitive systems process representations in order to produce particular behaviors.

However, here too, intentionality has its own problems. The main one has been labeled Brentano's Problem, and, like Kant's Puzzle, it is also a problem about causality. The problem turns around the cognitive capacity to misrepresent: I can have the false belief that outside is raining, even though it is a sunny day and act accordingly picking up my umbrella; I can suffer from perceptual hallucinations or perceptual errors. Misrepresentation, in some sense, is the capacity of representing a situation when such a situation is not real. In technical terms, misrepresentation is to have a representation that has no reference. The problem of misrepresentation shares with the problem of teleology that misrepresentation appears not to have a place in other areas of science. In the view of the world embraced by modern science of chains of causes and effects, there cannot be a break between two events. But it seems that misrepresentation precisely introduces such a hole or gap in the chain of events that produces behavior. If the behavior is explained by how a system perceives and responds to environmental conditions, then how is it possible that a system responds to the perception of an object that does not exist (as is the case with misrepresentations)? We cannot simply say that the mind produces a representation caused by some feature of the external world in order to produce the behavior because such feature does not exist. How is it possible that a representation be caused by a non-existent object? This is certainly impossible, isn't it? But we must nonetheless be able to explain how misrepresentations are possible, and this entails explaining how representations can go bad or wrong without assuming a gap in the causal explanation of cognition. Accounting for this issue is another central aim of this thesis.

The explanatory logic

The arguments in this thesis are entirely theoretical. I rely on different scientific insights and theories, but I shall not discuss them directly. Rather, I shall focus on the theoretical implications of a number of scientific theories in order to assess the

problems of teleology and of intentionality. In the spirit of Griesemer's words in the epigraph, my aim is to coordinate scientific models and phenomena in order to represent them in a particular way, and in relation to specific explanatory purposes.

That said, it is important to make explicit in advance the general explanatory logic I will apply in the analyses I will present here. My point of departure are a number of contemporary debates in the philosophy of biology. The state of art in this field shows that the mainstream view of biological theory —the Modern Synthesis— has been challenged by a fair number of scholars working in different fields in and out of biology. A common element in all these criticisms is the claim that developing organisms need to be put at the center of biological theory. I shall refer to this cluster of theories and research areas with the phrase *Developmental Turn*, which is an implicit acknowledgement that it does not yet constitute a coherent and robust theoretical framework. To be sure, my aim her is *not* to provide a comprehensive and well-articulated defense of the Developmental Turn, *or* to offer a detailed presentation of its challenges to the Modern Synthesis. Rather, my (more modest) aim is to explore some of the implications that the adoption of the Developmental Turn may have for the Life Sciences. The structure of this thesis is, therefore, that of a conditional statement: *If* we adopt the biological perspective raised by the Developmental Turn in the last decades, *then* what are the implications for the Life Sciences?

Specifically, I will focus on two main implications of the Developmental Turn. The first one has to do with the following question: what kind of theory of natural teleology should be defended *if* we accept the central tenets of the Developmental Turn? Or, in other words, what kind of solution to Kant's Puzzle may be offered from the perspective of the Developmental Turn? The second implication may be summarized thus: what would count as a solution to Brentano's Problem from the point of view offered by the Developmental Turn? Or, alternatively, how is intentionality to be naturalized *if* we adopt the general framework of the Developmental Turn? These are the main questions of this thesis and in the pages to follow I shall try to offer some answers to them.

Outline of chapters

The thesis is divided into three parts. Part I introduces in detail Kant's Puzzle and Brenatno's Problem and presents the mainstream frameworks that have been developed for dealing with each riddle: the Modern Synthesis in biology and etiological teleosemantics in cognitive science. This first part has three chapters.

In Chapter 1, I introduce Kant's Puzzle and Brentano's Problem. In both cases, I describe different wordings of these problems that have been around since they were first formulated. This will be useful to point out that the structure of both puzzles is the same: both teleological and intentional explanations are ex-

planatory necessary to understand biological and cognitive systems properly, but neither teleology nor intentionality can be easily accommodated in the foundations of modern science. Both problems are then about the tension that arises between what is explanatory useful and necessary and what is explanatory valid and legitimate in science.

In Chapter 2, I introduce what has been so far the mainstream view in biology: the so-called Modern Evolutionary Synthesis, forged during the first half of the 20$^{\text{th}}$ century. I will first focus on the forerunners of the Modern Synthesis and on the conceptual improvements developed by each of them. Next, I will present the main three pillars of the Modern Synthesis: (i) *Explanatory Externalism* —the idea that the only adaptive force in evolution is natural selection; (ii) *Replicator Biology* —the claim that inheritance is about the replication of the units of inheritance that carry all the necessary information to produce phenotypic traits; and (iii) *Populational causation* —the idea that evolutionary forces take place only at the populational level. I will conclude by pointing out three implications of these assumptions: first, that the core ingredients of natural selection —inheritance, variation, and fitness— are dissociated from developmental processes; second, that developing organisms are black boxed in the Modern Synthesis —they do not play any explanatory role; and finally, that the idea of evolutionary design naturalizes functional talk in biology, and consequently, teleological explanations too.

In Chapter 3, I introduce teleosemantics by first explaining what is the core of any teleosemantic proposal, namely the appeal to a notion of biological function in order to define the proper function of representational systems and, thereby, assess the problem of misrepresentation. I present the etiological theory of functions and its different supporters in order to point out how etiological teleosemantics integrates and is committed with the Modern Synthesis' framework.

Part II evaluates the adequacy of the mainstream answers presented in the first part of the thesis. As noted, my point of departure are the different debates in contemporary philosophy of biology and the need to re-think biological theory. These would be the grounds in which to assess the adequacy of the mainstream answers. This part is divided into two chapters.

Chapter 4 presents the *Developmental Turn*. My aim, as I already pointed out, is not to provide a defense of the Developmental Turn but an exposition of its main tenets and motivations for re-thinking biological theory. In this context, I will introduce three challenges to the Moderns Synthesis. Each challenge is connected to one of the Modern Synthesis' pillars: *Explanatory Internalism* and the crucial role of developing organisms are the two axes around which the critique of Explanatory Externalism is articulated; the *developmental conception of inheritance* is an alternative to the replicator view; and the statisticalist interpretation of natural selection challenges the very notion of populational forces. My conclusion will be

that, with the adoption of the framework of the Developmental Turn, the solution to Kant's Puzzle offered by the Modern Synthesis must to be rejected.

Chapter 5 extends the criticism of the Modern Synthesis to etiological teleosemantics. I present three challenges to the etiological theory of functions. First, I shall argue that the central role of development in evolution makes inadequate the idea that the presence of traits in evolution is exclusively a consequence of selection processes. Next, I show that etiological teleosemantics adopts many forms of Dichotomic Thinking that are rendered inadequate once the replicator view is abandoned. Finally, I argue that the criticism of statsticalism of the notion of populational forces is a direct attack on the foundational basis of etiological functions, given the fact that, under this interpretation of selection processes, etiological functions lose all their causal grounding.

Part III is devoted to the examination of possible alternative answers to Kant's Puzzle and Brentano's Problem. Following Eva Jablonka's advice in the epigraph, I meet the challenge and, after pointing out the weaknesses of the standard approaches, in this part I offer my attempt at providing a viable alternative. This last part has four chapters where I put together a number of proposals and ideas already present in he literature with my own personal take on the matter in order to articulate this alternative.

Chapter 6 introduces autonomous systems theory, which is the most comprehensive teleological theory available to date capable of disputing the mainstream position of the Modern Synthesis. I begin by identifying the historical roots of this theory, from Kant's own legacy to contemporary systems biology, through the organicist movement of the interwar period and the cybernetic school of the late 1940s and early 1950s. Next I focus on autonomous systems theory and on its account of autonomous agents, norms, goals, and functions in order to show that autonomous systems theory is a theory about the teleological character of physiological processes, in contraposition with the Modern Synthesis, which is about the 'teleonomic' character of evolutionary processes. Chapter 7 and Chapter 8 are about the missing element in this picture sor far, a process taking place at a different time scale from physiology and evolution: development.

Chapter 7 takes the first steps towards teleological account of development. Here I also start with an historical analysis, reviewing some important lessons to be learned from the old epigenesis vs. preformation debate, 19$^{\text{th}}$ century German teleomechanicism, and (developmental) organicism. On the basis of this historical analysis, I next introduce Denis Walsh's Agential Perspective, a teleological theory of development strongly influenced by the Developmental Turn. To conclude this chapter, I offer some remarks on Walsh's view and point out a number of shortcomings of his proposal. These motivate the alternative proposals presented in Chapter 8.

Chapter 8 is the core chapter of the thesis. Here I present and defend *Agential Teleosemantics*, a teleological view of development that is complementary to Walsh's theory. Agential Teleosemantics stands on two pillars: information processing and agentive capacities. I will argue that the informational processes regulated by developing agents result in adaptive ontogenesis and that this is a basic and necessary step for the naturalization of teleological development. Agential Teleosemantics states that teleological development involves *some sort* of intentional explanations, namely the idea that development proceeds by the normative uses of biological information. That is the main reason why I shall defend that development deserves to be treated in teleosemantics terms.

Chapter 9 extends Agential Teleosemantics to the cognitive domain in order to deal with Brentano's Problem. I open the chapter with a defense of the idea that development is the process where content is determined and norms are established. Thereafter, I move on to explain the difference between intentional explanations of development and intentional explanations in psychology. I close the chapter with some remarks on the different reasons and ways of extending the Developmental Turn into cognitive science.

Part I

Mainstream answers

Chapter 1

Kant's Puzzle on teleology and Brentano's Problem on intentionality

1.1 Kant's puzzle: you need teleology but you cannot naturalize it.

Teleology has always lived in a paradoxical atmosphere. It is an indispensable notion to understand, explain and interact with living beings. Yet, it certainly has not had a reputation worthy of being introduced in our scientific picture of nature. Kant's puzzle refers to this contradictory scenario. As a first approximation:

(KP) **Kant's Puzzle (rough definition)**
Teleology is reprehensible but inevitable in our understanding of nature.

1.1.1 What is teleology? First steps

Contrary to the folk conception, the word *teleology* was not introduced by Aristotle but by Christian Wolff (Gambarotto & Nahas, 2022; Van den Berg, 2013). Even though it is a concept that intends to refer to the natural world, there is not, for sure, a robust definition of teleology in the natural sciences; not at least as robust as there are definitions of other scientifically acceptable concepts. This is a hint that we should probably not be too eager to throw teleology out of the philosopher's office. The philosophers' attitudes notwithstanding, many biologists and historians deal with it as their first explanatory target. A central part of my project is to join these interdisciplinary efforts. Many other notions are related to teleology: purpose, goal, directness, intention, design, plan, etc. Part of my

analysis is devoted to finding the appropriate theoretical notions in order to better understand the experimental advances in biology that have been calling for the reinsertion of teleology in science. In order to take the first steps in this journey, let's get started with Kant.

Immanuel Kant (1724–1804) presented his view on teleology in the *Critique of Judgment* published in 1790 (here I will use the 2007 edition). It is, as expected, deeply integrated with all his philosophical enterprise, so, in part, it is necessary to understand the teleological flooring within the context of the whole building.

Kant correctly believed, and understandably appreciated, that organisms have a plus, something unique in nature, qualities that cannot be found in non-living systems. His worries about teleology, and their impact on post-Kantian biologists, are pretty much a consequence of the appreciation —and study— of this plus. Chiefly, this concerns the organizational and recursive properties of organisms. Each trait is connected with the whole organism and its functioning is dependent on that whole, while at the same time the functioning of the whole is dependent on each trait. The complex organization inherent to any living system brings to the fore many 'self-properties' (McLaughlin, 1990): organisms are *self-organized* wholes functioning by an ensemble of different parts guided by their inner and outer constraints and needs; the generation of this organized system is not brought about by an external source but by a process of *self-construction*, the material bases of which are internally produced, transformed by the organism itself; finally, the preservation of the organism —in its 'stable disequilibrium'— it is also work —metaphorically and literally, in its thermodynamic reading— done by the organism as a *self-maintained* unit. I shall return to these self-properties under their contemporary versions, where I will also introduce some new ones. By now the central point is that in all cases —organization, construction, and maintenance— the processes are guided by and according to organismal needs. These processes pursue a state that fulfills certain needs. For Kant, these kinds of processes deserve teleological explanations. The plus of organisms is related to the organismal needs and the kinds of properties involved in pursuing them. He called this plus *Naturzweck* (natural purpose) and it is the core of Kant's teleological thinking.

Kant believed that there is no scientific explanation for *Naturzwecke*. He thereby adopted an 'as much' strategy: you must avoid teleological explanations *as much as possible* in order to make science move forward while knowing that up to some point you will get stuck and inevitably rely on teleology. This means that *Naturzwecke* cannot be diluted: non-teleological explanations will never supersede teleological explanations. The main reason for such a position is his transcendental view of teleology. Explanations are epistemological artifacts. Kant's epistemological framework locates teleological explanations as a product of human understand-

ing (Desmond & Huneman, 2020). The 'Human Eye' has teleology in its retina and there is no scientific surgery capable of removing it.[1] It is usually considered that Kant's teleology plays a methodological and regulative role in science: it is not the source of a real scientific explanation, but it is just a guide to pursue scientific research (Lotfi, 2010). When a teleological notion is used in a methodological and pre-scientific way (not necessarily involving the whole Kantian framework), I shall call it *Folk Teleology*. Somehow the idea behind Folk Teleology is that there is something right in any teleological explanation: Folk Teleology catches some regularities of nature. However, teleological explanations cannot really account for these regularities in scientific terms, therefore the role of teleological explanations in scientific research is more regulatory or heuristic than truly explanatory.

Actually, as I will point out in the following subsection, there are many more problems around teleology, but the central one has to do with causation: teleology refers to the end-states of a system, and end-states cannot have causal powers on the system's current activity. Note that this is a *temporal* issue: how a system in time 2 (end-state) relates to the same system in time 1 (means towards the end). The problem of teleology is the *Temporal Problem of Causation*. Modern science, since Descartes and the Newtonian Paradigm, was built on the basis of the assumption that there are step-by-step causal interactions ending in the explanatory target. This view is usually considered to respect the *Causal Asymmetry Principle* of scientific explanations (Bromberger, 1966; Potochnik, 2017).[2] It roughly assumes that all scientific explanations are devoted to the understanding of the causal relations leading from time 1 to time 2, but not the other way around (that's why they are asymmetric). This is science, and this produces knowledge. Yet, teleology explains by focusing on end-states: somehow end-states in time 2 explain means in time 1. How is this possible? How can we predict, describe and comprehend an organism's life with an unfashioned explanatory strategy lacking scientific foundations? Why the 'Human eye' does not fit with the eye of modern science? Kant was clear about the impossibility to look at biology in the same way Newton dealt with physical phenomena:

> it is absurd [...] to hope that maybe another Newton may some day arise, to make intelligible to us even the genesis of but a blade of grass from natural laws that no design has ordered. Such insight we must absolutely deny to mankind. (Kant, 2007, 228)

[1] This is an extremely simple presentation of Kant's thought. It is nonetheless sufficient for my purposes here, since I shall not adopt a transcendental view of teleology. Later on, I will however attempt a more detailed discussion of the impact of Kant's ideas in contemporary thought (Gambarotto & Nahas, 2022; Huneman & Walsh, 2017).

[2] The principle is often referred to as 'Explanatory Asymmetry'. Insofar as it regards the asymmetry of causal explanations, I opt here for referring to it as 'Causal Asymmetry'.

The tension that arises once we try to deal with natural purposes is clearly expressed in *The Antinomy of the Teleological Power of Judgment*. As Denis Walsh stressed (Walsh, 2006b, 774), the puzzle is about the coexistence of the following thesis and antithesis:

> *The first maxim* of the power of judgment is the *thesis*: *All production of material things and their forms must be judged as possible on mere mechanical laws.*
>
> *The second maxim* is the *antithesis*: *Some products of material nature cannot be judged as possible on mere mechanical laws (that is, for judging them quite a different law of causality is required, namely, that of final causes).* (Kant, 2007, 214-215; emphasis in the original)

That is the core of Kant's Puzzle under Kant's view:

(KP) **Kant's Puzzle (Kant's view)**
Teleology cannot be avoided in the Human understanding of nature, even though it does not fit with the mechanistic explanations of science.

1.1.2 Setting the puzzle

This formulation won't be my target. My discussion will turn around Kant's puzzle but not under Kant's view. Two points need to be removed in order to achieve a more general and not theoretically committed presentation of Kant's puzzle: the transcendental view of teleology and the extension of teleology beyond organizational properties.

Concerning the former, if my attempt is to be part of a wave of neo-teleological thinking devoted to taking teleology as a genuine element of nature and reputable in scientific explanations, it cannot be considered just as an inevitable product of some Human Faculty —at least, not in a way different from the one other scientific notions are minted. It must be an intrinsic condition of life, so scientists can deal with it. The transcendental view should be replaced by an immanent view: organisms have a teleological dimension, and a coherent and complete scientific view of them cannot avoid such dimension. Contemporary scientific advances must have theoretical underpinnings capable of enabling scientists to say "look, that's teleology, no hidden mystery, it has a cause and it causes, and you can use it to explain, here is my paper".

The second point is a step forward towards a refined and global view of teleology. In part, this chain will not be complete until we reach the end of this thesis. So as Neurath's ship that is built as the journey proceeds, the meaning

and understanding around teleology would be constructed throughout the pages of this thesis. As explained, Kant was chiefly concerned with the organizational properties of organisms: how parts interact with one another in a complex, functional, and organized way. This is not to mean that he had considered teleology under this straight jacket; he certainly was also interested in generative issues. But as I will point out later, Kant's solution to the problem of causation cannot account for a complete teleological picture of nature, only for some aspects of it. Be that as it may, my aim is to define a comprehensive enough approach to Kant's puzzle that comprises other biological phenomena. I will distinguish two kinds of relationships that can be viewed as teleological, and, as teleology is a temporal issue, I will present three temporal levels of biological phenomena at which we can —or cannot— find teleology.

The two relationships that manifest the teleological character of an organism are (i) its fit with the environment, and (ii) the fit among its parts. As it can be appreciated, (i) refers to the interaction with the environment —the interactive dimension— and (ii) concerns the intrinsic organization —the organizational dimension; I will come back to this dichotomy in Chapter 6 and how it is presented in the contemporary literature. Surely, the notion of fit is not innocent in biology. It will take up many pages of this thesis. By now, it is enough to resort to a metaphorical notion: *harmony*. (i) and (ii) are harmonic in the sense that they provide functional and beneficial conditions for sustaining and producing life. If an organism does not interact harmonically with its environment, or if its parts have no harmonic relationships, its life decays. Sometimes it will be advisable to keep both properties, (i) and (ii), separated. But it will also be relevant to join them. The intersection point is the target of teleological explanations. I will refer to it as **aptness**: the adequacy of living beings to their life's conditions. So, I conclude, a global view about teleology must deal with aptness, encompassing the organizational and interactive fit of organisms. In other words, *teleology is there for explaining aptness*.

Turning now to the temporal dimensions, my discussion will be mostly based on Conrad Waddington's classification (Waddington, 1957, 6-7). He distinguished three different types of temporal change in biology:

Physiological scale: the organisms' physiological activities —e.g. metabolic, behavioral, etc.— at a particular temporal moment.

EXAMPLE: Bacteria use their flagella to move towards nutrient gradients and avoid toxins (a process known as *chemotaxis*). Mistakes could entail death —loss of metabolic resources or intoxication. How is it possible that a bacterium moves according to metabolic goals related to external environments and their future consequences? Another example: erythrocytes are produced in the bone marrow to be integrated into the circulatory system.

Different transportable elements reach different parts of the body thanks to erythrocytes, principally the distribution of oxygen by hemoglobin proteins. After a life cycle of about 120 days in humans, erythrocytes die and enter into a recycling process. In this lifespan, many organs and traits' functions critically depend on erythrocytes and, at the same time, these organs and traits also enable other traits to work. The point is quite clear: how is it possible that the production of erythrocytes carried out in one part of the body to act and live in another finally has consequences in others parts?

Developmental scale: the ontogeny of an organism —i.e. individual lifespan.

EXAMPLE: Dragonflies start their life in the ovum where organogenesis and morphogenesis begin. The second stage occurs in the water, as larvae, moving, eating, growing, and molting, until a certain stage, metamorphosis, takes place and the last molt gives rise to a dragonfly living outside the water. Each of these developmental stages, typical in insects, amphibians, and other taxa, are, for sure, highly complex. What is important to remark is that many of the capacities of further stages are inexorably dependent on the generation of traits in previous stages. How is it possible for a certain trait that is developed in a specific developmental stage —e.g. organogenesis— to have a crucial impact on further developmental stages —e.g. the larval stage?

Evolutionary scale: phylogenetics of populations —trans-generational time scale.

EXAMPLE: The fur of arctic wolves is particularly suitable for cold environment; birds' beaks are adequate for eating available foods; giraffes' necks allow them to reach higher branches full of leaves. How does a population evolve into a different one due to the consequences promoted by the evolved trait?

From here on, a number of strategies may be adopted: one could promote an explanation of the teleological dimensions at each level; or one could explain the teleological dimension of one level as a consequence of some feature present at another level, in such a way that the central target is the former and not the letter. One could even say that, after all, no teleology is needed and try to explain aptness in another way. All these options have been taken and I will present them throughout this thesis. Note that, moreover, there is an important difference in Waddington's classification: the first and second levels involve individuals; the third one involves populations. So one can argue that groups are the target of teleology and that this is the reason why there is a teleological dimension in the other temporal levels, or vice versa, that populational phenomena at the evolutionary

time scale manifest teleology because the members of the population are teleological systems. Again, these strategies are easily found in the history of biology and will be discussed later.

Before moving on, I would like to call attention to another crucial point: the populational and physiological level exhibit what I will call the $A-B-C-A$ *Phenomenon*. This is not the case of the developmental level, and the Gordian Knot about teleology, as I see it, is about this fact. In a nutshell, $A-B-C-A$ refers to the recursive character of a process. A causes B, B causes C, and C causes another A. At the physiological level, the reference to $A-B-C-A$ causal chains to approach teleological explanations has been usually adopted, at least since Kant, by the advocates of *organicism* (e.g., von Bertalanffy, 1969; E. S. Russell, 1945); see Etxeberria and Umerez (2006); Gilbert and Sarkar (2000); Nicholson and Gawne (2015) for historical introductions. For example, metabolic-based chemotaxis is enabled by the energetic resources of the cell, yet simultaneously the cell depends on its behavior to reach new energetic resources: behavior for metabolism and metabolism for behavior. The same can be said about the production of erythrocytes. They are needed for different physiological functions which, at the same time, are needed for the production and maintenance of erythrocytes. As we will see in the next chapter, at the evolutionary level, since Darwin, $A-B-C-A$ chains sit at the core of the *Modern Synthesis*' account of teleological explanations. The idea is that $A-B-C$ refers to those processes ongoing during an organism's lifespan which reproduces —if it is able to do so— and gives rise to a new organism (a new A). Successive transgenerational $A-B-C-A$ processes, mediated by reproduction and biased in an adaptive way —i.e. by natural selection, produce apt populations (more on organicism and the Modern Synthesis later).

However, development does not appear to present this property. In a nutshell, trivially, phenotypic outcomes do not feed back to those prior developmental stages that built them. The development of flies during metamorphosis does not depend on and is not casually connected to the consequences of a dragonfly using its wings. *At the developmental time scale, there is no recursivity between developmental stages* —understood in terms of $A-B-C-A$ chains as above. Is this problematic? It certainly is, especially for my naturalistic purposes. $A-B-C-A$ phenomena are the main source of understanding teleology in science. Yet development does not exhibit the $A-B-C-A$ property. So, what shall we do? As I see it, the first preliminary conclusion is that the Hard Problem of Teleology is a developmental one. To deal with and to offer a tentative solution to it will be my specific contribution to the understanding of teleology in this thesis —an issue, in my opinion, that has not yet been explored in detail.

Summarizing so far, I first argued for an immanent notion of teleology: as any other biological concept, teleology refers to the world of living beings. Next, I

underscored two relationships that endow organisms with teleology: the interactive and the organizational dimensions. They constitute the aptness of organisms, the main aim in any teleological explanation. Finally, based on Waddington's insights, I presented the three temporal levels at which biological processes take place, involving two kinds of entities: individuals and populations. Taking all this into account, we are in a position to give Kant's puzzle a more precise wording.

The problem can now be seen as a problem of naturalization: how can we naturalize the teleology present at each time scale? Naturalization is a common business in the philosophy and theory of sciences. For instance, many try to naturalize the mind so as to consider it part of nature and not a non-physical entity. This sort of project is primary metaphysical: to find the place of what was taken to be a non-natural entity in the order of nature. Note that this is not the main target of Kant's puzzle. The first reason for this is that metaphysical naturalization looks at itself as an oxymoron: to put it shortly, if one seeks to 'naturalize an entity', it is because the entity in question was already a natural one. That is, naturalization cannot take the form of the displacement of an entity from an ontological realm (a non-natural one) to another (a natural one).[3] The second, and more relevant, reason is that the naturalization of teleology pursues *epistemological naturalization*: how can we understand teleology as part of nature in a way that teleological explanations only appeal to the language of science? In this view, the project does not take a contradictory form: from the success (or failure) that we —philosophers and scientists— have in the project of naturalizing teleology, an impact will ensue in our epistemological resources, as individuals and communities, but not a change affecting those entities involved in teleological explanations. Epistemological naturalization is not about them, it is about us. I will henceforth use the term naturalization in this sense. Having said so, the formulation of Kant's Puzzle, the one that I will discuss at length, should take the following general form:

(KP) **Kant's Puzzle (General definition)**
Teleology is necessary to explain the aptness of biological phenomena at different temporal levels, even though it cannot be understood in naturalistic terms.

1.1.3 The Teleological Gap and three bridges

Kant's puzzle defines a gap between what is explanatory useful and what is explanatory admissible. As I explained, the core problem of teleology is backward causation. However there are other reasons that make teleology problematic which

[3] These reflections are inspired by similar musings found in Chomsky (2000).

arise when we connect it with other controversial notions in biological science: functions, norms, and intentionality. Intentionality is closely related to teleology, yet, while some teleological processes can be explained by invoking intentionality, it seems that it cannot be the source to explain aptness: while Humans do have intentionality, it seems that we cannot explain the physiology of digestive systems, chemotaxis in *E. coli*, metamorphosis or change in wolves's fur by attributing intentionality to those systems explained in teleological terms. Therefore, if teleology goes hand by hand with intentionality, we have the problem of ascribing intentionality to systems that, *prima facie*, do not posses intentionality; Part III of this thesis deals with this issue. Moreover, teleological explanation involves a normative dimension: if traits, developmental processes, or populations are directed to a certain purpose or goal, then there is a normative demarcation depending on whether the system achieves the goal or not. Notwithstanding, science is not supposed to be evaluative but descriptive. It is not about how things *must* be but how things used to be, are, and will be. Science should not be prescriptive but teleology introduces normativity in nature: more troubles. The problem of functions is indeed the same as that with normativity: the function of a trait (if defined by its purposes) introduces a demarcation of function and malfunction, a certain way in which a trait *must* function. All in all, this is the basic herd of nightmares that digs the gap between what is explanatory useful and what is explanatory admissible.

In this subsection I will review different possible strategies to bridge this gap, while I shall leave for the next subsection the presentation of a specific proposal. The central issue is: what is the place of teleology in the taxonomy of explanations? There are, *grosso modo*, three alternatives. Roughly, (i) to appeal to some form of reductive mechanism (*sensu* Nicholson, 2012), (ii) to fall back to vitalism, or (iii) to bring into play naturalized theories of teleology. At first sight, (iii) appears to be the most difficult position to maintain, but it is the one I will be arguing for here. Let's present them in detail in order to spell them out more clearly:

(i) Reduce teleology to mechanistic explanations and eliminate it from the vocabulary of science.

(ii) Assume the non-reducibiliy of teleology but give up any hope of integrating it in the vocabulary of science.

(iii) Assume the non-reducibility of teleology and introduce in the vocabulary of science.

Strategy (i)

The issue of mechanicism is extremely complex in philosophy of science. Part of the foundational issues around teleology turn around how we understand mechanisms —and how mechanicism is defended nowadays. I will now introduce three kinds of mechanicist thinking, following Nicholson (2012), and then I will present their connection with strategy (i).

Mechanicism: a phenomenon is explained by the interaction of the parts that integrate it.

Machine Mechanism: the use of a machine-like structure to explain a phenomenon

Causal Mechanism: a phenomenon is explained by a step-by-step causal chain that produces the phenomenon in question.

All forms of mechanicist thinking have traditionally been taken to be opposed to teleological talk. The reason is clear: the kinds of explanations they favor involve some form of forward causation, while teleology seems to appeal to backward causation. Eliminativist attitudes towards teleology take the form of some kind of reduction to the physicochemical level. Mechanicism puts parts in an explanatory prior position with respect to the whole: two classical proposals along these lines are those of Cummins (1975) and Craver (2007), for example. This entails that self-properties arise by the interaction of the parts, which, crucially, does not involve any intervening teleological notion. For example, cellular metabolism is understood as being the result of the action of many metabolites that, by themselves, do not exhibit teleology. If we can explain metabolism, along with its internal and external regulations, in this way, it seems that there is no need for teleology. A similar situation arises in the case of Machine mechanism. Importantly, the use of machines to understand biology is always idiosyncratic and strongly dependant on the kinds of machines available during a certain period in history. For instance, Kant, influenced by Descartes, took the clock as the machine-model. It is not too difficult to see that the work of the clock does not involve teleology.

Things get more interesting with Causal Mechanism. We can expect to identify some tensions here: step-by-step causal chains take the form of $A - B - C - D$ concatenated events, while in teleology (at least at the physiological and evolutionary levels, as organicism and the Modern Synthesis respectively suggested) we find $A - B - C - A$ chains. Is this an insurmountable obstacle? It is certainly the case that Causal Mechanism more or less explicitly endorses a form of the Causal Asymmetry Principle. It is about causal relations that produce a particular effect and the fact that no backward relation exists between causes and effects. In

this sense, Causal Mechanism may be compatible with some ideas that cannot be easily accounted for by Machine Mechanism or Mechanicism, such as top-down causation. In other words, Causal Mechanism is more permissive about the kinds of relations that may be taken as causal. We may therefore ask in which way a naturalized view of teleology can be made coherent with Causal Mechanism. This issue will be addressed in Part III; by now, it is enough to point out that the incompatibility between teleological and mechanistic explanations usually concerns Mechanicism and Machine Mechanism. The variety of theories adopting the principles of Mechanicism or Machine Mechanism though they "differ from one another about which causal principles are basic [...] they univocally reject explanations that appeal to vital forces and final causes" (Craver, 2013, 133–134).

Strategy (ii)

Strategy (ii), as I said, involves some form of vitalism: the idea that there is a non-material vital force causing the aptness of organisms (e.g., Bergson, 1907; Driesch, 1908). However, this is not really accurate. Vitalism —and its many variants— is only one of the theoretical stances that opts for strategy (ii). The proper label for this strategy should probably be *non-naturalism*: accepting that teleology is a real part of nature, but that it is not possible to explain it in scientific terms. For instance, Natural Theology, Divine Design or Intelligent Design are all proposals that posit a teleological dimension without —seriously— attempting to explain it from a scientific perspective. Among the advocates of this strategy we can find different positions, more or less metaphysically committed, some with only remote concerns with scientific aims, others with stronger methodological interests in scientific progress. Anyhow, these views have never been strongly defended within science. The reason is quite clear: they have no place in a naturalist picture of the world. Nonetheless, it is important to make explicit two points that will be retaken later: first, some versions of strategy (ii) were extremely fruitful in the attainment of scientific progress —both experimentally and theoretically; and second, many attempts at naturalization were deeply influenced by vitalist ideas.[4]

Strategy (iii)

This strategy aims at making teleological explanations scientifically tractable. Two options immediately come to mind: either we enlarge science, or we crop teleology. A first attempt might consist in adopting many insights of the vitalists while

[4] As we will see in Chapter 6, the foundations of organicism at the beginning of the 20$^{\text{th}}$ century were motivated, in part, by vitalist insights (Nicholson & Gawne, 2015). Moreover, as M. Weber (2022), Driesch's work in embryology, which pushed him to postulate a vital force, enabled many advances in the mechanistic understanding of epigenesis.

remaining strongly determined to tackle them with scientific weapons. This implies that there should be some area, discipline or theory in science capable of dealing with the teleological ideas usually associated with vitalism. We will see that this is in fact the case, particularly in the 21st century. The crop-strategy is to avoid any *mystical*, *non-natural* or *anthropomorphic* understanding of teleology —detach it from any vitalist reminiscence— and re-frame it in science without removing the scientific foundations. I call the first option the *top-down strategy*: trying to understand teleology in biology preserving some ideas of non-physical accounts. The second option is the *bottom-up strategy*: tackle teleology in biology by approaching it from physical-reductive accounts. Probably the most relevant naturalized theories of teleology take place at the intersection of these paths.

1.2 Brentano's Problem: you need intentionality but you cannot naturalize it.

Intentionality breathes the same paradoxical atmosphere as teleology. While intentionality is central in the explanation of behavior, we cannot explain it properly. Brentano's problem is about this tension. As a first approximation:

(BP) **Brentano's Problem (rough definition)**
Intentionality is reprehensible but inevitable in our understanding nature.

1.2.1 What is intentionality? First steps

As it is well known by everybody working in cognitive science, intentionality is the foodstuff of philosophers. Although many other aspects of cognition and animal life are treated by other sciences, intentionality has not yet been able to flee philosophy departments, at least not without the escort of a philosopher. Surely, intentionality is in itself a complex issue and most, if not all, cognitive science pivots on it. Yet, different scholars understand it differently. Many meanings have been attributed to and removed from intentionality: semantics, reference, meaning, intension (with an 's'), information, etc. I will discuss these terms and see which ones provide a promising avenue towards a solution to Brentano's Problem. As a first approximation, let's consider first Brentano's view on intentionality.

In his *Psychology from an Empirical Standpoint*,[5] Franz Brentano (1838–1917) appealed to the notion of intentionality in his aim to distinguish the mental from the physical. Since then, intentionality is taken to be *The Mark of the Mental* (or at least, one of its marks; Neander, 2017b). What traces the difference between mental and non-mental entities is that:

[5]The original edition is from 1874; I will use the 1995 edition.

> every mental phenomenon is characterized by what the Scholastics of the Middle Ages called the intentional (or mental) inexistence of an object, and what we might call, though not wholly unambiguously, reference to a content, direction toward an object [...], or immanent objectivity. Every mental phenomenon includes something as object within itself, although they do not all do so in the same way. In presentation something is presented, in judgement something is affirmed or denied, in love loved, in hate hated, in desire desired and so on. (Brentano, 1995, 68)

Although I will refine and update the notion of intentionality, this original formulation by Brentano is pretty adequate. It already incorporates two of the core constitutive ingredients of intentionality. The first one is *aboutness*: the idea that intentional states or processes are about or refer to something else; that they are directed towards an object. The intentional object that such processes or states are about is typically considered to be *content*.

The other notion is *normativity*, althoug it is not so easy to appreciate it in Brentano's quote. This will be clearer once the Cognitive Revolution has been introduced. The point turns around the *inexistence* of the object that an intentional state is about, that is, the fact that the intentional object has no reference in the world —there is no world-object but there is an intentional-object. Brentano's view relates primarily to the fact that, for example, we can think about things that do not exist, will never exist, or that may exist in the future. What this illustrates is that the relation between content and its reference needs not to be a real relationship; intentionality transcends what is actual. This suggests that, somehow, content and reference need not be causally linked. I will use the expression *Causal Mismatch* to refer to the possibility of detaching content and reference.

However, the main situation in which one detects the Causal Mismatch (the inexistence of the world-object) concerns in the *possibility of error* —also called "misrepresentation", an issue I will discuss at length here. As it will be clearer once I introduce the notion of representation, we can mistakenly represent the world. We can have hallucinations, perceptual errors, false beliefs. Errors clearly introduce a normative dimension. Here the connection between the Causal Mismatch and normativity is easy to appreciate: error is thr place to find the dissociation between reference and content; if this connection were a necessary one, then errors would never occur. Crucially, the possibility of uncoupling content and reference is central to explaining goal-directed behavior. The explanatory target is behavior as directed towards certain environmental conditions, as organisms are trying to fulfill some goals. Such goals introduce a normative dimension regardless of whether the goal is fulfilled or not. Therefore, that there is room for error is crucial for our understanding of goal-directed behavior. The possibility of error is a requisite for

normativity, and normativity is a requisite for goal-directed behavior.

Brentano argued that intentionality is explanatory and necessary to understand behavior. So we are again confronting an explanatory, useful and non-eliminable notion for the understanding of a natural phenomenon. We explain and interact with animal behavior as a phenomenon produced by how animals perceive the world and respond to it. We systematically explain behavior by positing a goal(normativity)-directed(aboutness) response. We will see that this is central to cognitive science. As it is widely recognized, explaining behavior by content is (i) explanatory useful in science —i.e. it allows to explain regularities and to predict outcomes, and (ii) it is systematic in our daily life and in the interaction and interpretation of the behavior of others. Traditionally, this explanatory strategy is called *Folk Psychology* —the use of psychological-intentional states to explain behavior. I prefer to use *Folk Intentionality* instead, in order to encompass a larger class of phenomena and to establish a parallelism with Folk Teleology. The central point is that Folk-Intentional explanations —as well as Folk-Teleological explanations— somehow require a solid scientific account. The reason is simple: such explanations are true (Dennett, 1987; Fodor, 1975). Or, at least, some truth is hidden behind them. If this were not so, why is Folk Intentionality so accurate at predicting, regulating and explaining behavior?

So far, I have not yet identified any problem connected with intentionality. There are hundreds of them. However, here, the core problem is causation, again. If the intentional states of the mind are about something that does not necessarily exist, it seems that we cannot explain behavior by content in causal terms: if behavior is caused by intentional states, and intentional states are caused by what they are about, but they can be about nothing, then there is a hole in the causal process that produces behavior. Something else must be going on in intentional explanations: the kind of causation at work in intentional explanations —if any— cannot be the one involved in other natural —and scientifically tractable— causal processes. If we combine aboutness and the normative dimension, we get stuck in a cul-de-sac. If behavior is explained by content, and there is no reference in the world for that content, there cannot be a causal chain behind intentional explanations.

Brentano's position concerning the naturalization of intentionality is quite similar to Kant's attitude concerning the naturalization of teleology: there cannot be a science of intentionality. Intentional causation is not within the scope of scientific inquiry. So we can understand Brentano's Problem as the problem of naturalizing intentionality. Intentional explanations are useful and work well in many cases, yet aboutness and normativity are intractable by modern science. This gives rise to Brentano's version of the problem.

(BP) **Brentano's Problem (Brentano's view)**

> Intentionality is necessary to explain behavior but science cannot explain the kind of causation involved in intentional explanations.

1.2.2 Spelling out the problem

As expected, Brentano's version of Brentano's Problem does not exactly match the themes of contemporary discussions of the subject. Intentionality came to the fore in science with the Cognitive Revolution in the mid 20th century. It was re-conceptualized and specified, and consequently awarded a central explanatory role in cognitive science. The problem of intentionality metamorphosed into the problem of representational states and their capacity to explain behavior reduces to the characterization of representational processes that produce behavior. Thus, representations and their manipulation became the core of intentionality in cognitive science. The content of intentional states is, therefore, the content of those representations that constitute the intentional state. Cognition is taken to consist of the manipulation of those of representations that lead to behavioral outcomes.

It is important to emphasize what is one of the main targets of mainstream views in cognitive science: the explanation of behavior. Of course, this is not an easy task and involves many different cognitive tasks, which somehow explains the interdisciplinary nature of cognitive science. A cognitive scientist —whether a neuroscientist, a psychologist, a linguist, a computer scientist, or an anthropologist— would probably need to integrate different kinds of knowledge from different disciplines and research areas to explain even the simplest behavioral phenomena. However, there is a central foundational issue in need of an answer: why is it not enough with neurophysiological explanations about the physicochemical processes in the brain and the body that produce a certain output? Why do we need to posit intentional states? Is it not enough with a causal description of neurophysiological processes? Cognitive science must have an answer. The central one is, as Dretske (1988) and Shea (2018) remarked, that such explanatory strategy would not account for the proper *explanandum*. Such strategy would not explain behavior as a consequence of an organism's goals (norms) directed (about) towards the environment. The outputs of cognitive processes are not just certain changes in the system under study, but those changes that are goal-directed. It is not just the output of a system, but *goal-directed behavior* a central explanatory aim of cognitive science. To deal with this *explanandum*, we need to introduce aboutness and the normative dimensions of goal-directed behavior, i.e. to tackle the *explanandum* we need intentionality. The emphasis on goal-directness is cardinal to argue for a cognitive enterprise.

Aboutness and the normativity of intentionality are easily recognizable from a representational viewpoint. Paradigmatic examples of intentional states are beliefs. John believes that the cat is on the mat. John's belief is about something: the

cat being on the mat. If he behaves in a certain way in relation with the cat, his behavior would be directed towards the content of those representational states that produce behavior. Normativity is also easy to see. A central phenomenon around behavioral errors is the phenomenon of misrepresentation. John went to the mat to catch the cat and screamed when he realized it was a rat. The success of John's behavior that fulfills his goal is partly related to accurate representation. As explained above, in the path towards naturalizing content one cannot dodge the issue of misrepresentations.

With this brief introduction to representationalism, let us now turn to see what form Brentano's Problem takes. The task is to understand aboutness and normativity on naturalistic grounds. The trouble is expected: if we explain behavior by content, and content needs not be based on a causal relationship with what it refers to, then it seems that intentional explanations cannot have a causal backup. The problem of causation requires a solution to the problem of content (what determines the content of a representation?) and the problem of misrepresentation (how is it possible to dissociate content from reference?). We can also present the scenario in the following way: neurophysiological causation produces organism's outputs. At this level, we do not need intentional terms. Yet we are not explaining behavior as the product of an organism's goal-states directed to the world. To introduce this level —which is where Folk Intentionality comes in handy— we need to attribute some contentful character to the neurophysiological processes. However, in the case of neuro-physiological causation, there is no problem of error insofar as the kind of causation involved is the mainstream one in science: interaction between neurons and different brain substrates that step-by-step produce behavior. There cannot be a gap in this process. But I already argued that at the intentional level the possibility of misrepresentation must exist. In this step, we lose the status of causation present at the neurophysiological level. Therefore, the issue around naturalization is how to preserve genuine causation without losing explanatory power. A non-existent cat cannot cause John's behaviour. Or can it?

After presenting the representationalist version of Brentano's Problem, let's call attention to a common strategy to circumvent the problem. We wish to explain how intentionality can have solid naturalistic underpinnings. For instance, we can ask where the content of the word *cat* comes from. One can answer that it came from our conceptual representation of a cat (the concept CAT). The word *cat*, let's suppose, derives its meaning from CAT. The point is that we can ask the same question concerning CAT. We certainly can answer that CAT takes its content from another representation (e.g. a perceptual image of a cat). Yet this strategy has to stop up to some point: you cannot explain the content of all representations by pointing out at the content of another representation. Sooner or later one has to reach intentionality coming from a non-intentional source. This is known as

underived intentionality or *original intentionality* (cf. Dennett, 1994; Haugeland, 1981). It will be the main target of Part III. What most naturalized theories of content seek is just to add a 'pinch' of naturalized intentionality. This is enough to season the rest of the dish. Of course, this doesn't mean that it is easy to derive the rest of representational contents in a system once you have this pinch (it is not just about compositionality). Notwithstanding, once you can explain this basic building block, you are showing that there certainly is nothing mysterious about intentionality. With these ingredients, the respresentationalist view of the problem may be formulated as follows:

(BP) **Brentano's Problem (Representationalist view)**
Intentional states, constituted by representations, are central to explain goal-directed behavior, caused by representational processes, but we cannot naturalize them.

Although in Part III I will discuss this particular formulation of the problem, I would like to present a broader, but still updated version of Brentano's Problem. The reasons are two.

First, Brentano's problem is about intentionality, not representations *per se*. As intentionality became to be understood in representationalist terms, then the target moved to representations. However, non-representationalist theories of intentionality in cognitive science —what I shall refer to as *radical post-cognitivism*— are gaining momentum and supporters. So I believe that a general definition of Brentano's Problem should be comprehensive enough to also encompass under this umbrella those radical views on intentionality. In this sense, Brentano's Problem arises both in representationalists and anti-representationalists theories of mind. This defines a common ground for discussion that facilitates the identification of influences and points of contact among different proposals. Even though I will not endorse a radical view, some of the proposals developed within this camp —usually strongly committed to naturalize intentionality— are linked with theories that will appear in this thesis (particularly, autonomous systems theory in Chapter 6 and ecological psychology in Chapter 7).

The second reason also involves a controversial issue in contemporary philosophy of the life sciences. It concerns the possible extension of paradigmatic animal properties well beyond animals, i.e. to organisms not possessing a central nervous system. There are, at least, three theses one may want to adhere to or reject. The first one is the Life-Mind Thesis: all living beings have a mind, and all mindful beings have life; life is coextensive with mind. A soft version of this is the Life-Cognition Thesis: life is coextensive with cognition. Adopting this thesis and rejecting the previous one entails that some cognitive phenomena are not mental (e.g. memory). One could even adopt a third strategy and claim that living beings

have neither cognition nor minds, but that they present agentive capacities. As it can be appreciated, the discussion partly hinges on how we understand the terms involved. So the question is not only about what organisms do, but also about how we define mind, cognition, and agency. I will not defend the first two options but I would support the third one. What is extended beyond the brain is not the mind but intentional states, therefore intentional explanations are also epistemologically central beyond animal activity. Be that as it may, understanding all living beings as agents will allow me to discuss all the aforementioned theses around the extension of paradigmatic animal properties. So a general presentation of Brentano's Problem must be related to how behavior —or any kind of motility (Dretske, 1988, Ch. 1)— is explained by the agent's inner states and their processing.

To conclude, here's the general formulation of Brentano's Problem:

(BP) **Brentano's Problem (General definition)**
Intentionality is central for explaining the goal-directed behavior produced by intrinsic properties of agents, even though we cannot naturalize intentional explanations.

1.2.3 The Intentional Gap and three bridges

Although cognitive science has constantly been producing a huge amount of data, experimental advances, technical improvements, and solid theories, all these scientific accomplishments are afflicted with the unsolved philosophical problem of intentionality. Beyond the nuclear issue of causation —involving a response to the content question and to misrepresentation— there are many, many other controversies. This is not only because intentionality is a tough nut to crack, but also because philosophy of language was the stellar philosophical discipline in the last century. Much ink has been spilled in order to analyze the phenomenon under different views and readings. The results were fruitful, such as the advances in logic, computational theory, or linguistics. Yet all these achievements were reached at the expense of creating new philosophical challenges that we are still dragging. Beyond causation, we can find in the following list some of the classical problems around intentionlity. The list is not meant to be exhaustive and one can easily get lost in this complex and mined territory:

1. The proximate vs. distal challenge: do we represent the distal object or the proximate stimulus?

2. The particular vs. general challenge: do we represent particular objects or the general class which they belong to?

3. 'Disjunctionitis' (Neander, 2017b, 149): which of the many properties of an object do we represent?

4. Frege's Puzzle, too: how co-extensive representations can have different contents?

5. The problem of non-actuality (already presented): how can there be contentful intentional states that have no reference?

6. The opaque context challenge: why representations cannot be substituted *salva veritate* in so-called opaque contexts? John believes that his flu is not dangerous but that Covid certainly is, even though his flu is Covid.

I will be mostly focused on the problem of intentional causation. As it is the case with teleology, the problem of causation in intentionality also defines a gap between what is explanatory useful —Folk Intentionality— and what is explanatory legitimate. Understood in causal terms, the question is: how can we bridge the gap between *flat* (neuropyshiological) causation and intentional causation? This gap could be also characterized in other ways: as the difference between syntactic (or correlational) information and semantic information or as the difference between natural meaning and non-natural meaning. In all cases, the first terms in these pairs are relatively uncontroversial for science —concerning its foundations. However, problems arise with the second terms of the pairs. Explaining them is a central *desideratum* in the foundations of cognitive science. How can we bridge this gap?

Here, also, we can identify three strategies to bridge the gap. The first is eliminativism, the second is dualism and the third one involves some kind of naturalized theory of intentionality. As expected, I will try to stay within the third group:

(i) Intentionality should be reduced to non-intentional explanations and eliminated form the vocabulary of science.

(ii) Intentionality cannot be reduced but intentional behavior cannot be correctly integrated into the vocabulary of science.

(iii) Assume the non-reducibility of intentionality and integrate it into the vocabulary of science.

Strategy (i)

This strategy adopts some sort of eliminativism. The clearest case is that of behaviorism (Ryle, 1949). Intentional states cannot be scientifically tractable for epistemological constraints: the mind is the land of the subjective and this is not territory for science. Therefore behaviorism is devoted to the study observable, objective patterns, such as stimuli and responses. Neuro-reductivist positions also

remove any reference to intentionality (Churchland, 1988). Although Folk Intentionality is heuristically useful, both in science and in life, all the genuine explanatory work is done by neural processes. Some *ascriptivist* theories (Shea, 2018) endorse an eliminativism of sorts (although this requires further discussion). Dennett's *Intentional Stance* (Dennett, 1987), for instance, argues that intentionality should be involved in explanations insofar as it helps in our understanding of science. Yet most scholars, but curiously not Dennett himself, understand this view as an instrumentalist position, where an intentional stance should be defended without any ontological commitments with particular physical units with intentional properties.

Strategy (ii)

I link this strategy with dualism: the idea that intentionality is detached (at least partially) from psychical phenomena and consequently cannot be tractable by science (e.g., Popper & Eccles, 1983). A paradigmatic example is Descartes. Yet, this position not only takes the form of Cartesian Dualism. Brentano also believed that intentionality falls outside the realm of science. Moreover, not all contemporary proposals in the philosophy of mind and language are committed with a naturalist position, or if they were, it would be difficult to anchor them within naturalism (as, for instance, Fodor and Pylyshyn (2015) argued at length concerning theories of concepts).

Strategy (iii)

Like teleology, the naturalization of intentionality also follows two paths: either we enlarge science or we crop intentionality. The first path involves preserving intentional explanations as first-citizens in science, arguing that modern science has the tools to deal with them: some scientific area, discipline, or methodology must exist to explain intentional behavior. The second option, the crop-strategy, consists of denying any non-natural *res cogitans* but taking intentional behavior as real and scientifically tractable. Here, I shall again call the first option the *top-down strategy*: approaching intentionality from dualist stances. The second option is the *bottom-up strategy*: approaching intentionality from physical-reductive accounts. The crossover of these two paths is a prominent naturalistic spot to rest.

1.3 Summary

It is perhaps not too pretentious to argue that intentionality and teleology occupy a similar place in the philosophy and history of the life sciences. I have been explicitly pointing out some connections between them, but also implicitly by following

the same pattern in their presentation. There are good reasons for claiming that this sameness is not a coincidence but rather a consequence of the intrinsic relation that exists between intentional and teleological phenomena. This thesis is about this relationship. In this introductory chapter I focused on presenting a brief and relatively shallow overview of the central *desiderata* and their motivations, and on presenting different categorizations and terminological issues that will be recurrent throughout these pages. Figure 1.1 on the following page summarizes the connections between intentionality and teleology. It is about their riddles. We will see how they are also connected when it comes to find solutions. The next chapter presents the most famous solution to Kant's Puzzle: Neo-Darwinism. Chapter 3 in its turn, deals with the most recurrent solution to Brentano's Problem: teleosemantics. These are the mainstream views in the business. Let's sink our teeth into them.

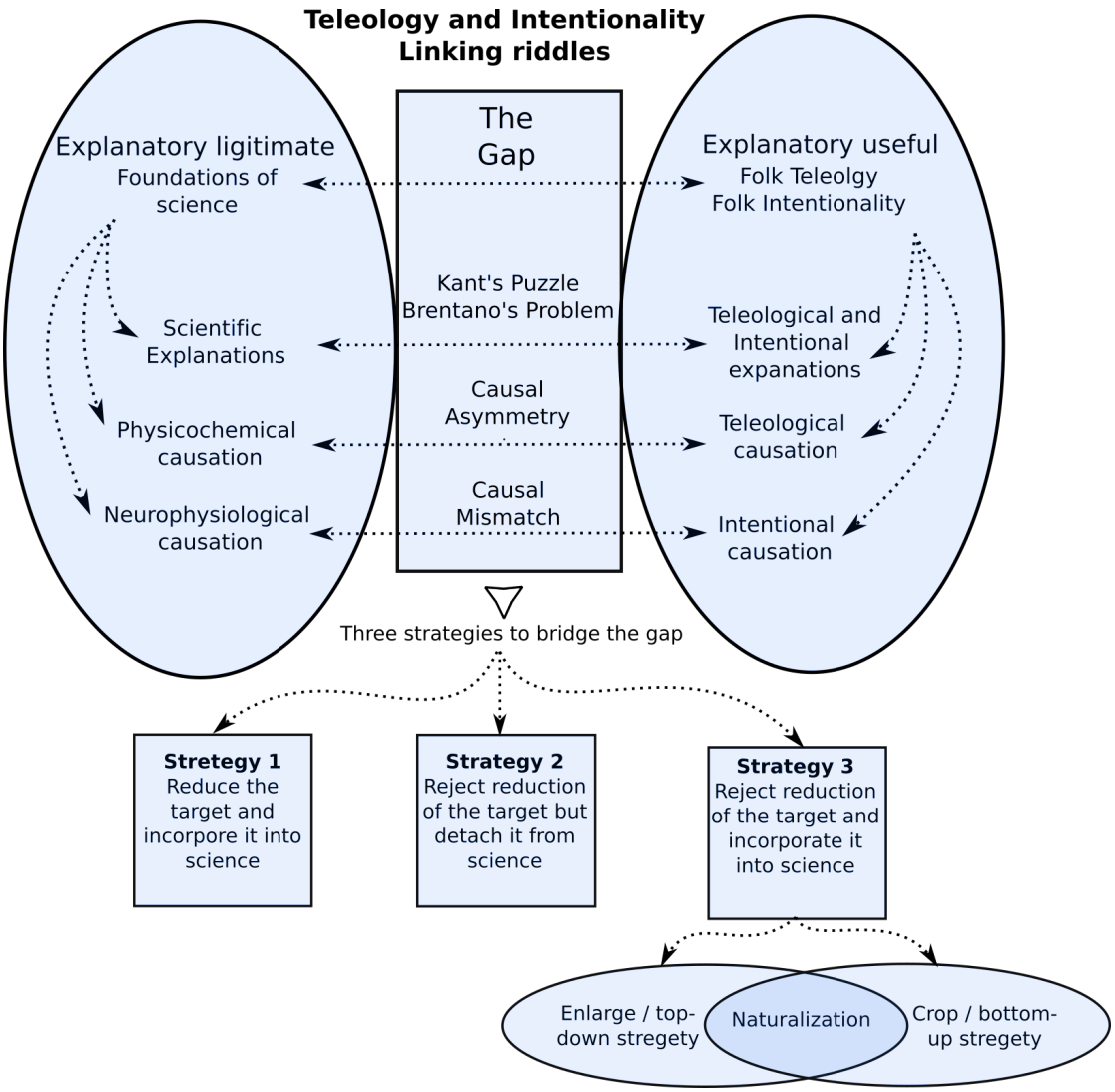

Figure 1.1: Teleology and Intentionality: linking riddles; (adapted from Rama, in press).

Chapter 2

The Modern Synthesis on teleology

> The theory of descent alone can explain the developmental history of organisms.
>
> Ernst Haeckel, 1866, 7

2.1 The gestation of the Modern Synthesis

Throughout this section, I will be introducing important figures in the history of evolutionary thought with the aim of presenting the mainstream view in evolutionary theory forged in the first half of the twenty century: The Modern Synthesis (henceforth MS).

2.1.1 Darwin

Charles Darwin (1809–1882) carefully read Paley's book on Natural Theology. Indeed, in a letter to his friend John Lubbock, Darwin wrote: "I do not think I ever admired a book more than Paley's Natural Theology: I could almost formerly have said it by heart" (quoted in Ågren, 2021, 19). Despite Paley's influence on Darwin's thought, and his motivation around Paley's interests, Darwin was not so keen on Paley's answers. Darwin's theory of natural selection could therefore be presented as an alternative, non-Paleyan solution to Paleyan problems (Ruse, 2003, 2019). Paley's focus of attention was the design-like character of organisms. Definitively, in the light of our considerations in Chapter 1, this is a striking fact once one thinks about it. The parts of an organism constitute a coherent whole,

in such a way that each part performs certain functions that fit both with the rest of the organism and with its environmental conditions, allowing living beings to grow, live and reproduce. Paley's answer was theological; whence its position outside the scientific picture drawn by modern science since Galileo and Descartes. This is also the reason why Darwin was not satisfied with Paley's view. According to Paley, just as the design-like character of a clock —the machine metaphor of his times— implies the existence of a watchmaker, the design-like features of organisms is to be taken as evidence of the existence of a designer; an intelligent designer: God. Darwin's inspiration was Paley's emphasis on the importance of explaining the design-like character of living beings, which seems not to be present in non-living (and non-externally designed) things. His answer is well known: his theory of natural selection presented in *The Origin of Species* (Darwin, 1859). As John Maynard Smith once stated: "The main task of any evolutionary theory is to explain adaptive complexity, i.e. to explain the same set of facts which Paley used as evidence of a Creator" (Maynard Smith, 1969).

A central insight in Darwin's theory is the significance of populational and historical explanations to tackle the two main *explananda* in biology (Sterelny & Griffiths, 1999; Walsh, 2003): adaptation —the design-like character of living beings— and diversity —the *design space* that species can explore, i.e. the possible repertoire of phenotypes based on the scope and constraints on variation (Dennett, 1995). Natural selection is a process that relies on three variables: variation, fitness, and inheritance. Firstly, to select something (a trait), you need more than one option (variation is needed). Secondly, there must be a reason for selection (differences in fitness). Finally, if the bias introduced by selection is not preserved through time, (i) we would not find stable groups (species) of selected options (traits), but rather an ever-changing state of affairs; and (ii) selection processes would be one-shot trials (i.e. operating within one generation). So inheritance is needed to preserve the bias and give natural selection enough time to do its work.

Darwin's main emphasis was on fitness (inspired by Malthus's *An Essay on the Principle of Populations*). The key was not only that fitness introduces a bias in nature, but that this is an *adaptive* bias. Here the central three insights of Darwin's proposal come to the fore: the *struggle for life*, its consequences in *populational explanations*, and the production of *adaptations*. 'Struggle for life' is usually interpreted in the sense of competition for survival and reproduction between members of the same species. Malthus argued in his book that individuals compete whenever the resources necessary for life are limited. Those who strive are able to reproduce, such that better-adapted individuals will spread and outdo those who lose the contest. The expression 'struggle for life' need not be taken so literally, however. In particular, it may involve more than just competition. For Darwin, adaptive evolution is a consequence of individuals struggling

for life. Struggling for life should encompass all the individual processes that determine the fitness of an organism. Fitness differences among organisms —the raw material of selection— are a consequence of the aptness of each individual, and what individuals do to survive and reproduce. The real adversary of life is death, not conspecifics; struggling for life is better represented as a fight for life and reproduction.

Crucially, as noted before, Darwin's breakthrough was to perceive the importance of a populational level of analysis to understand the fit and diversity found in nature. In order to see how differences in individual fitness produce complex taxonomies of adapted groups, we need a historical perspective. We can see, therefore, natural selection operating through time (thanks to inheritance), and producing different trends biased by fitness differences. Populational explanations thereby place the question at the proper level of analysis and connect individual fitness differences with the diversification of adapted populations. This Darwinian insight was labeled *Population thinking* by Mayr (1975), which he opposed to pre-Darwinian *Typological thinking* (but see Amundson, 2005). Importantly, if selection is a mechanism that introduces an adaptive bias through time, the outputs of this process should explain the design-like character of living beings. Such outputs are *adaptations* (although adaptation is also used to refer to the very processes of natural selection; Sober, 1984). Individuals become adapted to their living conditions and differentiated from others by natural selection operating on heritable variations on fitness between individuals throughout history.

In Darwin's theory —and this remains an accepted view within the MS, the bias of selection is pictured as a slow and gradual process, acting on successive generations to accumulate inherited variations. This view of selection is known as *gradualism*, and it opposes *saltationist* views, such as that one defended by Goldschmidt (1982) and, later on, by Gould (1989). Importantly, both saltationist and gradualist views are nonetheless based on natural selection and its central ingredients. So, as we will see regarding other assumptions of evolutionary theory, gradualism is not a prerequisite for evolution by natural selection.

While fitness was at the core of Darwin's theory, he did not possess a robust account of phenotypic variation; he also explicitly acknowledged the want for an adequate theory of inheritance (while accepting a Lamarckism of sorts). This fact teaches us something interesting. The structure of natural selection is indeed quite irrelevant to the mechanisms of inheritance and variation. As we will see, under the umbrella of the three ingredients of natural selection (variation, inheritance, and fitness), there have been many experimental and theoretical proposals concerning how they should be understood and explained. Darwin's theory occupies a specific place under this umbrella. I will presently show how the neo-Darwinian Modern Synthesis occupies another one.

2.1.2 Mendel, Weismann, Morgan, Johannsen

Gregor Mendel (1822–1884) carried out his experiments during the second half of the 19$^{\text{th}}$ century (Mendel, 1866), but they were only rediscovered and appreciated in the first years of the 20$^{\text{th}}$ century. They were soon perceived as crucial to overcome the difficulties confronted by Darwin's theory. To be sure, Mendel's model seemed to contain the essentials to elaborate an adequate theory of inheritance for natural selection —even though the suitability of Mendel's inheritance for Darwin's natural selection was not immediately appreciated. The experiments carried out by Mendel concerned variations in the traits of plants, and just indirectly on variation in inherited materials. The crucial insights are twofold. First, phenotypic differences underscore heritable differences. Mendel's experiments on how traits are obtained by reproduction connected phenotypic differences in parents with phenotypic differences in their offspring, in a way that whatever material differences are transmitted from parents to offspring, they explain the similarities of traits among generations. Second, such heritable phenotypic differences are due to stable and discrete units of inheritance. Stability allows natural selection to trust the inheritance mechanism. If inheritance mechanisms were not reliable, natural selection would not be capable of constructing stable clusters of similar phenotypes. Moreover, as units of inheritance are linked to discrete phenotypic differences —what Mendel appreciated in his experiments, inheritance is also understood as the transmission from parents to offspring of discrete and differentiated units. Mendel certainly did not have a robust and complete theory about the mechanism and materiality of such units of inheritance —indeed, only recently have we been able to articulate a well-articulated, albeit incomplete, theory of inheritance (cf. Section 4.2). Notwithstanding, Mendel's experiments align well with natural selection: variation on fitness as a consequence of phenotypic differences is connected with variation on inheritable units, in a way that, by selecting traits, natural selection spreads the units of inheritance.

August Weismann (1834–1914) sowed one of the central ideas in evolutionary thought: the idea that the germline and the somatic line are independent of each other and do not interact (Weismann, 1892). This gave rise to what has come to be known as the *Weismann barrier*. As we will see later, Weismann's view represents the first movement towards the separation of development and inheritance. He located inheritance at the germline (what is usually known as *hard-inheritance*), firstly by positing that germ cells are sequestered at the beginning of development —they remain unchanged during development— and secondly by arguing that at least some changes in somatic lines are not transferred to other generations. In a nutshell, this implies that developmental processes cannot affect inherited material —thus blocking the possibility of the Lamarckian inheritance of

acquired characters.[1] Selection picks up the fittest traits, as Darwin showed. Such traits are connected with stable and discrete heritable units, as Mendel defended. Weismann's theory represented a further step forward: heritable units needed for evolution should be somewhere on the germline.

Thomas Hunt Morgan (1866–1945) and his Mendelian Chromosomal Theory of Inheritance puts even more emphasis on specifying the location of heritable units of variation needed for natural selection (Morgan, Sturtevant, Muller, & Bridges, 1915). Such units are within germ cells, in the chromosomes. The explanatory logic remains intact: natural selection operates with heritable units of variation producing fitness differences. The location of such units within the cell makes explicit a necessary commitment to make natural selection work. It concerns the Genotype-Phenotype map (GxP), i.e. development. If selection operates over phenotypes, but it distributes heritable units located within the cell (chromosomes), the connection between such units and the selected phenotype must be robust to see fitness differences as a consequence of variation on heritable units. If this were not the case, the reason (fitness) why some traits are selected would not be transferred down to the next generation, and thus evolution would not take place. There must be a robust connection between chromosomes and phenotypes if selection is going to produce adaptive complexity over time.

These units of inheritance were eventually given the name of *genes*. In 1909, Wilhelm Johannsen (1857–1927) introduced the demarcation between the genotype and the phenotype. This is connected both with Mendel and Weismann's works and later on with Morgan's chromosomal theory. Mendel's studies on inheritance were based on the idea that the remembrance of phenotypic products obtained by the transmission of units of inheritance —somewhere within plants' seeds. Moreover, Johannsen's distinction is also connected with the Weismann barrier and the separation between germline and somatic line. It is in this context that Johannsen introduced the word *gene*. The first solid genetic theory was that of Mendel and known as Mendelian genetics (Griffiths & Stotz, 2013, ch. 1). The Mendelian gene is indirectly predicted on the basis of phenotypic outcomes. Lenny Moss recently introduced the distinction between Gene-D(evelopment) and Gene-P(henotype) (Moss, 2003). While the former concerns the mechanistic role of genes in development, the latter concerns the connection between the gene and the phenotypic products —what Mendelian genes refer to. Mendel's genes must be somewhere in the seeds of the plants he used in his experiments.

The genetic talk was central in advancing inheritance theory; Mendel, Weismann, and Morgan did much to fix up Darwin's weakness. However, genes were

[1] However, as Amundson (2005) argues, the rejection of Lamarck's ideas was not Weismann's central aim but it was just a consequence of trying to provide a coherent connection between inheritance and evolution.

also crucial in another improvement to Darwin's theory: population genetics. As it can be noted, the material basis of genes —as well as the mechanism of inheritance— is not an indispensable element to do evolutionary biology. Indeed, population genetics was developed before the discovery of the *material* gene (cf. section 2.2). As already noted, the fact that different positions can be taken within the umbrella of natural selection illustrates that natural selection could be considered as a general explanatory frame rather than a unique way of understanding evolution.

2.1.3 Fisher, Wright and Haldane

Karl Pearson was the *pater familias* of the biometric school by the end of the 19th century. But the integration of mathematical tools to model populational processes in biology was forged by three central figures during the first decades of the 20th century: Ronald A. Fisher (1890–1962), John B. S. Haldane (1892–1964), and Sewall Wright (1889–1988). As it is well known, there was an intense debate between these scholars, but their work was central for evolutionary theory to take two important steps forward. First, they provided a robust mathematical framework for the kinds of populational explanations required to understand natural selection processes. Darwin's claim concerning the importance of populational explanations eventually found an answer and its grounding under the form of a set of mathematical tools capable of describing the dynamics of populational changes through time. The second big step concerns abstraction. This will be a key issue in this thesis. The mathematical underpinnings of populational genetics made it possible to deal with populational properties independently of what is going on at the individual level. Abstraction became present in all the ingredients for natural selection:

Inheritance: A key concept also introduced in the 1930s is that of *heritability*, which quantifies what proportion of variance of a particular trait in a particular population is genetic variance. Crucially (broad) heritability is not about *how* parents transfer traits to their offspring, but it simply expresses a ratio between genetic and total variance of a specific trait in a specific population (Keller, 2010). Therefore, heritability is independent of the mechanism of inheritance. Moreover, it is a property of populations, not of individuals. Heritability necessarily requires taking groups as the units of analysis to reach the statistical measures concerning trait transmission (Godfrey-Smith, 2009).

Fitness: As noted before, fitness plays a central explanatory role in the attempts to understand evolution as an adaptive process. However, while Darwinian

fitness applies at the level of the individual, to the individual processes of the struggle for life, population genetics is based on the notion of trait fitness, a statistical measure of the fitness values within a population (Ariew & Lewontin, 2004; Matthen & Ariew, 2002). Trait fitness, but not Darwinian fitness, is the central explanatory concept in populational biology (cf. section 4.3 for discussion). Although there are other important issues concerning the technical notion of fitness in the history of evolutionary thought, it is the difference between Darwinian (individual) fitness and trait (populational) fitness that represents a quantum leap of evolutionary theory in its journey towards a understanding of natural evolution.

Variation: Variation was also treated at the populational level in the mathematical models of the biometric school, although here lie some of the tensions between the different proposals. As expected, the source of variability was posited at the genetic level. Crucially, what the term 'gene' refers to and how genetic variation operates in ontogeny are independent issues. The material basis of genetics is dissociated from the logic of populational explanations. The relevant point is that whatever genes are, they must comply with the requisites posited by Mendel (discreteness and stability), Weismann (sequestration and hard-inheritance), and Morgan (transparent GxP map). Consequently, while rejecting many possible sources of heritable variation, the biometric school proposed three principal ways of genetic variation within a population: mutation, migration, and drift.

The compendium of the ideas briefly presented in this section was pictured as a synthesis, the so-called *Modern Synthesis in Evolutionary Biology*. It is not easy to present the MS, either as a biological theory or from a historical perspective. There is no consensus concerning the birth of the MS, although its basic building blocks were put together by different *architects*, such as Theodosius Dobzhansky (Dobzhansky, 1937), Ernst Mayr (Mayr, 1999), Julian Huxley (Huxley, 1942), and George Gaylord Simpson (Simpson, 1944). Moreover, the MS cannot be construed as a set of axioms or natural laws for biology, but rather as a particular way to understand natural evolution, consisting of different insights that were being integrated as the process of building a coherent framework to look into nature went on. If something unifies the MS, it is a particular stance concerning the explanatory logic of natural selection. Within this logic —presented as a synthesis of the ideas sketched in this section— different empirical advances were made possible. In the following section, I will introduce three pillars of the MS and how the synthesis was later consolidated by the discovery of the material *molecular gene*.

2.2 The Pillars of the Modern Synthesis

2.2.1 Explanatory Externalism

Explanatory Externalism is the thesis that the only directive force in evolution is natural selection (Godfrey-Smith, 1996; Walsh, 2015).[2] This does not necessarily entail that natural selection is the unique factor in evolutionary processes but that it is the only one that introduces an adaptive dimension. As explained, selection processes act over many generations, operating on heritable variations underlying different fitness values, and leading populations to an adaptive port or towards extinction. The process of selection is therefore the only factor that allows treating the output of populational processes in adaptive terms. This idea became standardized at the core of the MS and, for example, it was openly formulated by Huxley in a letter to Mayr: "*Natural selection*, acting on the heritable variation provided by the mutations and recombination of a Mendelian genetic constitution, is the *main agency* of biological evolution" (quoted in Huneman, 2017, 71; emphasis in the original).

Externalism hence refers to the fact that the adaptive bias in evolution comes from outside of the organism; from the environmental conditions that organisms confront and that determine their fitness values. Although it seems paradoxical, if we want to explain Paley's concern about the adaptive complexity of organisms, organisms are explanatory indispensable. Externalism stands in opposition to internalism. According to the latter, internal organismal processes also introduce an adaptive element in natural evolution. Ever since Darwin, many disputes in the history of biology have turned around the externalism vs. internalism debate.

The externalism vs. internalism debate will reappear later on. For now, we have enough information to understand why the MS did not treat phenotypic variation as an adaptive process, thus putting all the explanatory burden outside the organism. This idea is a consequence of the aforementioned view of inheritance and variation forged by the precursors of the MS. In a nutshell, insofar as developing organisms do not constitute a real source of heritable variation, the internal processes undergoing adaptive changes during ontogeny would not participate in evolutionary processes. In other words, natural selection is blind to internal adaptive changes led by the organism itself. Even though such adaptive changes —as we will see presently in connection with such processes as niche construction, phenotypic plasticity, and self-organization— do in fact alter the fitness values of individuals, they cannot enter into the evolutionary scene because they

[2] There are other externalist accounts not based on natural selection. Paley's theological view is also externalist, for example. Moreover, there are also externalist positions not based on the populational/historical level. As I will discuss in Part III, McShea (2012) and Babcock and McShea (2021), for example, propose an externalist account based on field theory.

are not inherited. So, as I will analyze in detail later on (Section 4.1.2), the MS took two explanatory strategies to deal with internalism. First, to claim that adaptive developmental changes are just a consequence of prior selection processes on hereditary units (genes). Second, to argue that the role of developmental processes in evolution is not a relevant force of evolution because only hereditary units (genes) play the role of specifying developmental outcomes.

The impact of these two strategies in biology will be discussed later on, in Chapter 4, both from a diachronic and from a synchronic perspective. The two strategies block internalism. The explanatory burden lies on natural selection acting on hereditary units. It is therefore crucial for the MS to maintain that hereditary units are not part of developmental processes. While pre-Weismannian theories of inheritance were epigeneticist (Amundson, 2005), that is, they conceived of inheritance as the transmission of developmental resources, inheritance and development became dissociated (Nicholson, 2014; Walsh, 2007a, 2013c) due to Weismann's Barrier and Morgan's view on the GxP map. As I will argue in the following subsection, Explanatory Externalism depends on a theory of inheritance that is divorced from a theory of development. This view has come to be known as Replicator Biology.

But before getting into that, there is a residual issue that needs to be solved. If phenotypic variation is not an adaptive force in evolution, what kind of force is it? We already pointed at one ingredient of the MS's answer to this question: the issue is not about phenotypic variation, but about variation at the level of hereditary units (genes). So, how did the MS understand genetic variation? Notoriously, the MS understood genetic variation as a random process. Randomness is a source of debate and confusion, because 'random' does not mean that we cannot know the causes of variation. It is not a limit to what is knowledgeable. It also does not mean that all variations are equally probable —as is the case when we flip a coin. Randomness rahter refers to the non-adaptively directed character of genetic variation (Griffiths & Stotz, 2013; Sober, 1984). Genetic variation does not take place anticipating the adaptive consequences it might have on adult phenotypes —whence many biologists's use of the term *blind* instead of random to avoid confusion and make explicit that variation is blind to its consequences. Variations are detached from their adaptive implications. Of course, this does not entail that variations are not adaptive, but that their existence is not connected with such adaptive consequences.

Let's pause for a moment and think about teleology for a while. As stated in the previous chapter, teleological explanations concern the adaptive character of living beings. Explanatory Externalism locates teleological explanations outside the organism (more on this later). Since phenotypic variation is not considered to be an adaptive force, it cannot be the source of teleology. Another way to

appreciate this is to focus on the randomness of variation. Teleological explanations concern the adaptive consequences of a certain phenomenon, in a way that the phenomenon can be explained by such consequences; but, since variation is random, teleological explanations have no place here: variation does not occur anticipating its adaptive consequences.

The picture that emerges from Explanatory Externalism is that of selective pressures posing problems and natural selection choosing among many random phenotypic variations. Problems are the environmental conditions that organisms confront. Selection processes lead to adaptive evolution by choosing the better solutions that arise randomly. It is a weird game: to solve a problem, you can not know it in advance. Random variations arise without knowledge of the environmental conditions they will contribute to overcome. Populations have to wait for a solution that fits with the problem. If the solution arises by an adaptively directed process of phenotypic variation —such as developmental plasticity— this solution is considered not to be reached according to the rules of the game, because it would not be inherited. Evolution only plays with random solutions. The only non-random process is that of selection.

As Lewontin notoriously argued (Lewontin, 1983b), this view encourages a peculiar conception of ecological niches, where niches are seen as preexisting environmental conditions that organisms must confront. Niches are not defined in relation to the organisms's properties, but as riddles that genetic variation must find a solution to. This is clear in the idea of adaptation. As Lewontin illustrates, an adaptation fits with a preexisting condition. If I travel to a different country, I may need an adapter to plug and charge my computer. I need something that makes the output of my computer fit with the inlets of that country. I need an adaptor. Explanatory Externalism entails the *Autonomy of the Environment*: the idea that environmental conditions are independent of organisms. Explanatory Externalism will be challenged in Section 4.1. Now, let's move to Replicator Biology.

2.2.2 Replicator Biology

Organisms die. Evolution continues. Organisms are mortal, but those things that evolve are not. Organisms appear to be sitting at the wrong level of analysis to speak of evolution. This idea is central in Replicator Biology. While those who brought Replicator Biology to the first page of the life sciences were George C. Williams and Richard Dawkins (Dawkins, 1976, 1982; G. C. Williams, 1966), its main tenets had been around in biology for some time already. The relevance of the units of evolution was first pointed out by Weismann along with his remarks on the continuity of the germplasm. As already noted, the construction of stable and immortal (in the sense that they transcend a single lifespan) units of inheritance

was a task taken up by many scholars besides Weismann. As Dawkins himself acknowledged, he is in debt with many biologists that contributed to the foundations of the idea of the replicatior, such as Ronald Fisher, John Maynard Smith, and of course, George Williams (Ågren, 2021, 12).

I already noted how the hereditary units of evolution, reduced to genes, are central in the foundations of the MS. The reduction of evolutionary processes to the genetic level was accomplished without really knowing the physical substrate of genes. The central framework that gave rise to replicator ideas was already present well before even knowing the molecular structure and functioning of genes:

> The key take-home message is that the gene's-eye view wants to talk about genes in an abstract way and happily accepts a bit of fuzziness regarding their physical basis. (Ågren, 2021, 52)

The post-Darwinian view on the units of inheritance was reinforced with the discovery of the structure of DNA by Watson and Crick in 1953 (Watson & Crick, 1953), what came to be called the *molecular gene* (Griffiths & Stotz, 2013). Mendel's atomism, the Weismann barrier, and Weismann's ideas on sequestration and hard inheritance were reformulated in molecular terms —i.e. they were re-expressed as a 'molecular Weismannism' (Baedke, 2018a, 30), in such a way that Morgan's demarcation between inheritance and development became accepted and experimentally supported. This connection is explicit in the Central Dogma of Molecular Biology presented by Crick (1958, 1970). It concerns the flow of information in the expression of proteins during development. It is about how evolutionary adaptations genetically inherited unfold to construct the trait that has been selected. As the mantra of the Central Dogma reiterates once and again, the flow of information is from DNA to RNA and from RNA to proteins acting on cells. This is how organisms are built. Information flows from DNA to cells, but crucially not the other way around.

As expected, the causal primacy in development relies on the inherited unit: DNA. There is a crucial theoretical reason why the Central Dogma needs to be defended within the MS framework. If it is posited that inheritance relies exclusively on amino acid sequences, then such sequences must bear a strong, causal correlation with the traits produced. That is what the Central Dogma secures. The Genotype-Phenotype map is transparent enough for us to see the selection of traits as related to inheritance units. If the Genotype-Phenotype map were opaque, and the direct correlation between genes and phenotypes got lost, the explanatory strategy of the MS, based on a gene-centric view of evolution, would fail. Once inheritance is reduced to DNA sequences, the Central Dogma must be defended. As noted, the molecular gene had reinforced erstwhile replicator ideas. Williams and Dawkins came to make them public and explicitly central in evolutionary thought. Before moving to the contemporary presentation of replication

due to Richard Dawkins, let's first introduce another key piece in order to fully set the scene: Erwin Schrödinger's *What is Life?*

Schrödinger's famous book was the main motivation for the construction of molecular biology, and the work of Watson and Crick was a clear consequence of this influential book. There, Schrödinger openly sets out the question concerning the origin of adaptive complexity. How is it possible that such complex, organized, and highly functional systems (with their distinctive thermodynamical properties) have come to existence out of non-living matter? He provided two possible answers: (i) Order-from-order and (ii) Order-from-disorder. Option (i) was the one advocated by the Central Dogma and the molecular revolution in biology —beanbag genetics, as Mayr called it. The idea is that the organization and complexity of organisms arises from a non-living organized system, a kind of aperiodic crystal. The order present in living beings comes from an already ordered entity. Such entity eventually came to be the molecular gene. Option (ii), closer to organicism, posits that the distinctive qualities of living beings emerge from the wholeness of organisms. They are not to be found or explained by looking at the sub-organismal level; for an overview of Schrödinger's views and their later impact in 20$^{\text{th}}$ biology, see Walsh (2015, ch. 1).

The path in the history of biology traced till now anchors Dawkins' view. The importance of Schrödinger's question and the molecular gene as an answer to it is connected with the shift from an organism-centerd view to a gene-centered one scaffolded by the MS and Replicator Biology. Now Biology is not about organisms but about replicable, sub-organismal, complex, organized units bearing the history of evolution encrypted in a molecular code. Biology can live without organisms; genetic talk encompasses all biology.

The gene-centered view of Replicator Biology entails the distinction between replicators and vehicles. Replicators have three properties: (i) longevity, (ii) copy-fidelity and (iii) fecundity (Dawkins, 1978). On the one hand, replicators preserve with high fidelity the features of previous generations in a way that their life transcends a single lifespan. On the other hand, vehicles are the places where replicators reside. They just live for one generation, and their distinctive features are not reliably present in further generations unless they are caused by replicators. In Dawkins's view, genes are replicators and organisms are vehicles. Vehicles carry replicators, the latter transform the properties of the former into a suitable format to confront selection pressures: phenotypes.

There is a clear division of explanatory labor between replicators and vehicles in evolutionary processes. Replicators are the leading characters. They store the information acquired during evolution and transmit it down to the next generations. Replicators are responsible for producing the complexity of organized living beings.

Replicators are central in the MS view of evolution; Replicator Biology is above all a particular way of understanding inheritance. It fits with, and it was indeed motivated by, the precursors of the MS. Replicators build copies of organisms, organisms confront selective pressures. How such copies are made is completely irrelevant. Vehicles are relevant only because they are the link between selection and inheritance. Inheritance is about the replication of sub-organismal units responsible for producing a copy of the parental vehicle. The process of producing copies is secondary; the role of vehicles in evolution is subordinated to that of replicators. Dawkins himself stressed that the notion of *vehicle* was proposed "not to praise it, but to bury it"(Dawkins, 1994, 617). The division between replicators and vehicles and their different explanatory roles in evolution lies at the core of the MS. It is, after all, the demarcation of inheritance and development. Let's conclude by illustrating this view with the following quote from Dobzhansky, one of the MS's architects:

> Heredity is, in the last analysis, self-reproduction. The units of heredity, and hence of self-reproduction, are corpuscles of macromolecular dimensions, called genes. The chief, if not the only, function of every gene is to build a copy of itself out of the food materials; the organism, in a sense, is a *byproduct* of this process of gene self-synthesis. (Dobzhansky, 1958, 21; emphasis added —quoted in Jablonka & Lamb, 2020, 3)

2.2.3 Populational causation

The picture of natural selection as a populational force was formulated and popularized by Elliot Sober in his *The Nature of Selection* (Sober, 1984), but similar versions of this view had been previously proposed by other MS biologists. Sober proposed to see populations as entities and, establishing a parallelism with the Newtonian paradigm, that different forces act on such entities. Forces cause objects to move. The momentum of a car is the sum of all the forces impinging on it. Similarly, the movement of a population (that is, the changes in populational structure) that gives rise to new adaptations and species is also the result of many forces acting on it.

According to Explanatory Externalism, natural selection is an evolutionary force, but, although it is the only *adaptive* force, it is not the only *force*. Mutation, migration, and drift are all cases of population-level phenomena that provide the causes of evolution. Crucially, the assertion that evolutionary forces are population-level phenomena promotes the preclusion of individual-level phenomena in the explanation of evolutionary processes. As I already remarked, Darwin's central explanatory insight was the role of populational processes to explain adap-

tive complexity and the diversity of life. In Section 4.3, I offer an analysis of how the MS populational thinking is different from Darwin's populational thinking. For the time being, the following quote should suffice to have a picture, even if a rough one, of the core assumptions of populational causation:

> The population is an entity, subject to its own forces, and obeying its own laws. The details concerning the individuals who are parts of this whole are pretty much irrelevant…In this important sense, population thinking involves *ignoring individuals*. (Sober, 1980, 370, emphasis in the original)

The theory of forces proposed by Sober is connected with his view on evolutionary functions. This idea was also defended by others before, such as Maynard Smith (1978, 23), who stated that "the 'function' of an organ is taken to mean those of its effects which have been responsible for its evolution by natural selection" (quoted in Lloyd & Gould, 2017, 51).[3] This is an assumption that lies at the core of etiological teleosemantics, as we will see in the next chapter. Sober distinguishes between selection-for effects and selection-of effects. Both effects are defined at the populational level and concern evolutionary processes. However, the central difference regards causation. On the one hand, selection-for effects define those functions that a trait-type performs and that have contributed to that trait being selected. They concern the causal role that a trait has had during natural evolution: "When there is selection for one trait and selection against another, the traits *make a causal difference* in survival and reproductive success" (Sober, 2013, 339; emphasis in the original). On the other hand, selection-of effects do not contribute to the fitness values in selection processes. They are just by-products or consequences of a trait being selected-for. Selection-of effects are causally epiphenomenal in evolution and causally dependent on selection-for effects.

This distinction can be illustrated with the following example: Imagine someone has a salt shaker with two kinds of salt. One is thin and pink, and the other is thick and white. When seasoning up one's dish, only the thin and pink grains of salt will go through the holes of the salt shaker, while the thick and white ones will remain inside. In this scenario, thin salt is selected for seasoning up the food because being thin is what allows it to pass through the holes; its whiteness is not a force because in the processes of selection this property does not play any relevant function, it was only selected of the population of grains of salt in the shaker.

[3]More recently, Alex Rosenberg, vindicating Wright (1976), has also declared that: "It was Darwin's achievement to show how functions arise and persist in the biological domain by showing that they are all of them adaptations. Biological functions are naturally selected effects" (Rosenberg, 2020, 17).

As it can be expected, the connection between selection-for effects (as causal phenomena underpinning selection processes) and Explanatory Externalism lies in the fact that selection-for effects are the causes of adaptations. As noted, forces move populations. If adaptive complexity is caused by external selection processes, adaptations should be the result of the forces acting on populations. As Sober explains:

> A is an adaptation for task T in a population P if and only if A became prevalent in P because there was selection for A, where the selective advantage of A was due to the fact that A helped perform task T. (Sober, 1984, 208)

The location of the causes of adaptations at the evolutionary level was also defended by Ernst Mayr through the distinction he introduced between ultimate and proximate causes (Mayr, 1961, 1974), a distinction that is also central to his view of adaptive complexity and the design-like character of organisms. Mayr's proposal is well known and has been amply discussed: ultimate causes concern evolutionary processes taking place during the history of populations; proximate causes lie at the individual level and concern those physiological and developmental processes that take place during an individual's lifespan. This distinction between kinds of causes is also connected to a division of explanatory labor. Some areas in biology are devoted to proximate causes —physiology, morphology, developmental biology— while others explain by positing ultimate causes —evolutionary biology. The responsibility for dealing with the former resides in functional biologists, while evolutionary biologists are in charge of the latter.

As expected, Mayr's explanation of the design-like character of organisms relies on ultimate causation. Those explanations involving ultimate causation are responsible for explaining what a trait is for and why it is present in nature. This is directly connected with his account of teleonomy. Mayr borrowed this term from Pittendrigh (1958) to explain the design-like character of organisms in order to avoid the connotations usually associated to the word 'teleology'. By staying away from interpretations involving non-natural forces, his view aspires to be scientifically reputable. This idea is explicit in his commitment to 'mechanistic purposiveness' (Mayr, 1961, 1504): the teleological character of organisms may be accounted for in terms of purely mechanical explanations.

A teleonomic system is a system that has been programmed and whose activities are carried out according to such a program. This includes machines, insofar as machines are controlled by a designed program. Living beings, according to Mayr, also obey a program: a genetic program. Organisms are teleonomic because they are genetically programmed (Mayr, 1974).[4] What organisms do during

[4]Monod and Jacob (1961) and Jacob (1993) also developed influential teleonomic views of

their lifespan —those processes involving proximate causation— is a consequence of their inherited genetic program. But, unlike machines, genetic programs have not been designed by an engineer with a mind. The responsible for programming genes is natural selection, the process in which ultimate causes act. The view is then quite transparent: natural selection designs programs, organisms obey natural selection by being genetically programmed. Since organisms are programmed, organisms are teleonomic systems. Even though Mayr is clear about there not being any teleological or teleonomic character in natural selection, the causes of the organisms's teleonomy are ultimate causes. Note, finally, that purposiveness is mechanistic and derives from the fact that the implementation of the genetic program is mechanistic —that is, the process of protein expression, cell formation, and other individual-level processes guided by genes.

The view of natural selection as a causal process is connected with the debate between the positive and the negative interpretations of natural selection (Walsh, 2000). I will discuss this in detail in Chapter 4, but I shall offer a brief sketch of it here. The positive reading of natural selection posits that natural selection is a creative force. It brings about new and adaptive organisms. Contrary to this view, the negative one contends that natural selection is just a filter. Natural selection needs pre-existent variation in order to select, but the process of selection does not bring anything new into the world. It just distributes or eliminates what already exists.

As expected, most MS biologists support the positive view (cf. Beatty, 2016, 2019, for a detailed historical analysis). Natural selection causally contributes to the existence of adaptive complexity. Natural selection creates the complex adaptive systems that we find in nature by being an adaptive force guiding populations to adaptive scenarios. In Mayr's words: "It [natural selection] acts as a positive force that pays a premium for any contribution toward an improvement, however small. For this reason, profound thinkers about evolution, such as Theodosius Dobzhansky, Julian Huxley, and G. G. Simpson have called selection 'creative' " (Mayr, 1988, 45-46). Stephen Gould also put the creativity of natural selection at the core of (neo-)Darwinism:

> Why was natural selection compared to a composer by Dobzhansky; to a poet by Simpson; to a sculptor by Mayr, and to, of all people, Mr. Shakespeare by Julian Huxley? I won't defend the choice of metaphors, but I will uphold the intent, namely, to illustrate the essence of Darwinism —the creativity of natural selection. (Gould, 1992, 44)

I will conclude this section with an important historical remark, connecting evolutionary causation with ideas already introduced above. The emphasis on

adaptive complexity based on the idea of genetic programs designed by evolutionary process (cf. Keller, 2002, for a critical analysis).

populational causes pushed developing organisms outside of evolutionary biology, as I will discuss in the following section. The role of individual causation was put aside during the gestation of the MS. I previously argued that the MS was able to avoid taking phenotypic variation as an adaptive force in evolution. There were two strategies for doing so: reducing phenotypic variation to the genetic level and stressing the causal primacy of genes in development. In Chapter 4, I will argue that these strategies were applied to specific proposals during the first half of the 20th century. In general terms, this phenomenon can be understood as a process of theoretical reinterpretation. In particular, we can find in the history of biology different notions that used to refer to the individual-ontogenetic level, and then became *reinterpreted* in populational/evolutionary terms (Amundson, 2005).

Godfrey-Smith explains how Darwinian natural selection was reframed by the MS: "A noticeable difference between Darwin's descriptions of natural selection and most modern summaries [...] is that the recent ones do not refer to a 'struggle for life' " (Godfrey-Smith, 2009, 48). As explained before, the struggle for life refers to those processes that cause Darwinian/individual fitness. In the MS framework, the causes of evolution are populational, and the relevant fitness is not Darwinian but populational (Matthen & Ariew, 2002). The MS defends a different view based on abstraction. As many definitions of natural selection exemplify (e.g. Lewontin, 1970, 1; Levins & Lewontin, 1985, 76; Ridley, 2003, 71-72), variation, inheritance and fitness are abstracted away from individual-level processes and defined at the populational level.

2.3 The legacy of the Modern Synthesis

While modern evolutionary theory was being developed, a number of theoretical commitments were endorsed that are *not* a prerequisite for natural evolution. That is, one could abandon them without actually denying evolution. Of course, experience is the backup for theoretical discussion —in an ideal philosophical world. So once I have presented the experimental advances —such as the molecular gene and the mathematical theory of populational biology— connected with the gestation of the theoretical underpinnings of the MS (the pillars of the MS), let me consider now some of the consequences that the erection of such pillars entailed.

2.3.1 Dissociated processes

The fragmentation of biological processes at different levels has been remarked by many (e.g. Amundson, 2005; Keller, 2010; Walsh, 2007a, among others) and already highlighted in this chapter. During the gestation of the MS, a particular position was adopted concerning central biological phenomena: inheritance was di-

vorced from development. While inheritance was introduced in biology (from social science) to refer to the process of phenotypic construction based on the transfer of developmental resources from parents to offspring, thanks to Weismann and Morgan, inheritance and development became two different episodes in the living world. Inheritance occurs prior to development, at conception (Mameli, 2005). Evolutionary relevant (heritable) variations occur before the manifestation of such variations at the phenotypic level. Finally, the emphasis on populational causes and on natural selection as the only adaptive force in evolution has relocated fitness at the populational level —contrary to Darwin's original suggestion. Individual development is put aside in evolutionary explanations based on fitness values. As I will present in the following subsection, the divorce between the ingredients of natural selection and development not only entailed the exclusion of developmental processes from evolutionary theory, but also the explanatory vacuity of the very organism. All the ingredients of natural selection can be understood without reference to organisms and focusing only on supra-organismal and sub-organismal phenomena (Nicholson, 2014; Walsh, 2015).

It is important to note that the fragmentation of biological processes is not a requisite for natural evolution. Darwin's theory, for instance, based on a Lamarckian theory of inheritance, did not separate individual development from inheritance. Moreover, throughout the first half of the 20th century —the period that gave birth to the MS— were put forth a number of different proposals about phenotypic variation arising from developmental processes, and not from a genetic source. Classical examples are James Mark Baldwin's 'new factor in evolution' (later come to be known as the Baldwin Effect), Conrad Waddington's genetic assimilation, or Ivan Schmalhausen's stabilizing selection. Importantly, their proposals perfectly fit in and are aligned with a theory of natural selection. That is, the sources of phenotypic variation proposed by these scholars were assumed to be evolutionary relevant in the processes of natural selection.

2.3.2 Black-boxing developing organisms

"Something very curious and interesting has happened to biology in recent years. Organisms have disappeared as the fundamental units of life. In their place we now have genes, which have taken over all the basic properties that used to characterize living organisms [...] Better organisms made by better genes are the survivors in the lottery of life" (Goodwin, 1994, 1).

Brian Goodwin starts his influential book by pointing out this striking phenomenon: biologists dispense with any reference to the organism in order to explain organisms. While the *explanandum* concerns organisms, the *explanans* refers to genes. Organisms became explanatory irrelevant in evolutionary theory. Insofar as evolutionary theory plays a unifying role in biological theory in the MS,

and it provides the explanations of adaptive complexity, organisms were displaced as second-class citizens in biological theory and in the understanding of adaptive complexity. The externalist framework of the MS just looks at the relationship between inherited inputs (genes) and selected outcomes (traits) without actually paying attention to how such a relation is constructed throughout ontogenesis; philosopher of biology Ingo Brigandt summarizes this state of affairs as follows: "a selection-based explanation of phenotypic evolution merely requires *that* genetic differences result in phenotypic differences (so that variation is heritable), and it is irrelevant *how* genetic differences developmentally lead to phenotypic differences" Brigandt (2013, 84, emphasis in the original).

Organisms are unique. That much was already appreciated by Kant. However, instead of looking directly at those properties that are present in organisms and not in non-living beings, the MS put the emphasis on the sub-organismal and the supra-organismal levels. Populational changes in molecular units explain life. Inherited genetic variations are distributed in populations according to their fitness values. As a result, we obtain adaptive genetic programs that are responsible for producing organisms and their unique qualities. Within this scenario, it is expected that organisms are treated as "*vehicles* in which replicators travel about" (Dawkins, 1982, 82; emphasis in the original), as "merely the *medium* by which the external forces of the environment confront the internal forces that produce variation" (Levins & Lewontin, 1985, 88; emphasis added), as the "*arena* in which this interaction [genome variations and natural selection] is played" (Michel & Moore, 1995, 127; emphasis added), as "mere *middlemen* in evolution, a sort of *interface* between the organism building activities of replicators and the selecting role of the environment" (Walsh, 2006b, 775; emphasis added), or as "the *superficial face* that genes show to the world" (Sober, 1984, 228; emphasis added).

The preclusion of organisms is connected, as Viktor Hamburger emphasized in his contribution to Mayr and Provine's *Evolutionary Synthesis*, with the so-called *black-boxing* of development (Hamburger, 1980). The explanatory vacuity of development is connected to the aforementioned fragmentation of biological phenomena. None of the relevant ingredients needed for natural evolution concerns individual lifespans. What took place during development remains only for that organism. If development were a relevant source of phenotypic variation, or if there were epigenetic systems of inheritance acting during ontogenesis, then development could not be black-boxed. But we saw how and why, according to the MS, this is not the case. Development can be ignored because, as was stated by Haeckel (1866, 7) long ago, "the theory of descent alone can explain the developmental history of organisms".

The intrinsic and unique properties of organisms were ignored by evolutionary theory, and their explanatory role was neglected by adaptive evolution. Similarly,

development was absent from the process of natural selection. Development was disconnected from fitness, inheritance, and variation. I could sum up this by saying that *developing organisms* (Baedke, 2018b) were black-boxed by the MS. Even though the main target of biology is to understand what is inside the box, one can explain adaptive complexity and unify biological theory without actually looking into it.

2.3.3 Evolutionary design: Solving Kant's Puzzle

We have enough information already to see how the MS solved Kant's Puzzle. Darwin was the key. As Mayr (1998, 131) once said, "Darwin had solved Kant's great riddle". Moreover, we know the specific proposal defended by the MS within the umbrella defined by natural selection. That is, we know the form that evolutionary theory took and its relation to the explanation of adaptive complexity (that is, the target of Kant's Puzzle). So let's take stock and sum up all these ideas.

As far as certain details are concerned, the solution to Kant's Puzzle provided by Darwin is not so different from the one suggested by Paley. In both cases, the source of teleology is extrinsic. Extrinsic teleology is in opposition to intrinsic teleology. While the former posits that the processes responsible for explaining the design-like character of organisms and their harmony with their conditions of life are external to organisms, the latter defends that those specific properties of organisms, not present in non-living beings, underlay teleological explanation. In the third part of this thesis, I will defend an intrinsic view of teleology. Both Paley and Darwin posited an external source of teleology. In both cases, such an external source allows us to introduce teleological vocabulary in nature under the possibility of considering organisms as designed systems. Of course, the difference lies in the kind of source posited. Paley's view is theological, Darwin's is scientific. Paley's God is replaced by Darwin's natural selection.

Any naturalist solution to Kant's Puzzle must be aligned with the Causal Asymmetry Principle (1.1.2). Darwin's view on adaptive complexity seems to fit well with such a principle. Accordingly, the goal or purpose of a trait is connected with the adaptive role of that trait during evolution. Those functions that a trait performs during evolution make it an adaptation. This is a historical process. Past events involving selection, inheritance, and variation explain adaptive complexity. The goals and purposes of the trait's functions are defined by successful selection processes in the past. In this sense, teleological explanations can be about goals without involving any kind of backward causation. Past causes are posited in teleological explanations to define the teleological functions of traits. The results of such evolutionary processes are evolved and adapted systems: "Organisms are adapted, hence they are teleological, and (for the Darwinian) this teleology can

be explained through, and only through, natural selection" (Ruse, 2000, 223). As noted before, this solution involves $A--B--C--A$ chains. The process that introduces the backward loop is that of inheritance. In other words, without a reliable inheritance mechanism, $A--B--C--A$ chains would be broken: it would not be possible to build a new A that preserves many of the properties of previous generations. By biasing these chains via the process of selection — the unique directive force— we can see, first, how evolution is endowed with a directive dimension towards the production of adaptive complexity, and why there is no backward causation involved in evolutionary processes.

In sum, Darwin made it possible to speak about the design of organisms. Even though there are many discussions within Darwinism about the metaphoric character of teleological language, the explanation of the teleological dimension of living beings hinges on everything that falls under the umbrella term of *designed by natural selection* (Ayala, 2007; Dennett, 1995; Gardner, 2009; Ruse, 2000, 2003). Surely, many biologists adopt an eliminativist position of teleology based on Darwin's theory of natural selection (e.g. Godfrey-Smith, 2009; Rosenberg, 2011, 2014). So there are disputes concerning whether Darwin´s theory is a naturalist solution to Kant´s puzzle or whether it is just an eliminativist account of teleology (Veit, 2021). But those that sympathize with non-eliminativist stances see natural selection as the process responsible for introducing teleological language in nature. Even, as Lennox (1993) argues, Darwin himself saw natural selection as a way of legitimizing teleological and functional talk in biology. Moreover, Darwin was explicitly worried by the use of the word *selection*. He intended to separate his view of natural design from any proposal involving intentional design, such as Paley's view. His doubts concerning selection had nothing to do with its inadequacy for dealing with adaptive complexity, but with the fact that he just wanted to make clear that selection is not an intentional process. This may be taken to be another plus of Darwin's theory: the explanation of teleology does not presuppose any previous goal-directed system/process/entity, so circularity is avoided.

The MS view falls under Darwin's view, but, as already remarked, it proposes a specific stance to understand the process of natural selection. Here the pillars of the MS come to the fore. To the extent that natural selection is the only driving force in evolution, and adaptive complexity is explained in evolutionary terms, teleological explanation falls within the logic of Explanatory Externalism. An external factor defines the purposefulness and goal-directedness of traits. The process of reproduction, central in the recursive character of selection processes necessary to account for the Causal Asymmetry Principle, is understood as replication. Finally, the causal underpinnings of teleology lie at the populational level, in populational causation; i.e. in populational forces operating during history, 'from the past to the present'.

Clearly, not all MS biologists would accept the teleological side of their theory. In part, this is a consequence of the negative connotations teleology has had in the history of biology usually attributed to vitalists. However, if we keep in mind that a naturalist theory of teleology does not involve any non-natural entity or tension with the foundations of modern science, but it is just connected with the explanation of the adaptive complexity of living beings, then we can link the view of the MS on adaptive complexity with the design-like character of living beings. Of course, words are relevant. That's why, for instance, Mayr adopted a different terminology; others use teleological language (purposes, goals, design) just as metaphors. Be that as it may, the MS and its Darwinian underpinnings provide a clear and robust answer to Kant's worries. Organisms are internally and externally apt due to natural selection.

2.4 Summary

In this chapter, I introduced the mainstream solution to Kant's Puzzle: the Modern Evolutionary Synthesis. I started with Darwin and his theory of natural selection. However, the maturation of evolutionary thought that gave rise to the MS involved a particular way of understanding natural selection. As already noted, it is relevant to realize that the form that evolutionary thought took is not the only possible one. In other words, rejecting many of the MS ideas does not turn one into an anti-evolutionists. Under the blanket idea of evolution by natural selection, there is room for many possible positions to choose concerning the three basic ingredients: inheritance, fitness, and variation. Even though neo-Darwinism took many insights from Darwin, in many respects, neo-Darwinism is not Darwinian, in the sense that Darwin accepted many ideas that are considered mistaken by the MS. As Jablonka and Lamb recently commented:

> Mayr and the other subscribers to the MS thus excluded from it certain theories (Lamarckism, orthogenesis), outlooks (essentialism) and mechanisms (soft inheritance). There was no room in the MS for any non-gradual, goal-directed or internally driven processes, and no room for the inheritance of acquired characters or any other type of 'soft inheritance'. Darwinism was redefined: 'The term "Darwinism" in the following discussions refers to the theory that selection is the only direction-giving factor in evolution' (Mayr, 1980, 3). This was certainly not Darwin's Darwinism — it was a version of neo-Darwinism, but labelling this view as 'Darwinism' undoubtedly endowed it with more authority. (Jablonka & Lamb, 2020, 8)

There are two important technical and theoretical advances of 20th century

biology that have helped the MS build its basic pillars. One is the introduction of mathematical tools to understand evolutionary processes. This improvement gave rise to populational genetics, a research area at the supra-organismal level. Another is the discovery of DNA, which contributed to the consolidation of Weismann's and Morgan's views. The field of molecular biology, a discipline at the sub-organismal level, fitted well with with the MS view on inheritance —as replication, variation —as a random process, and development —as process controlled by inherited units. In this journey, developing organisms eventually became explanatory vacuous. The next chapter is about teleosemantics. We'll have to wait until Chapter 4 to analyze the adequacy of the MS from a contemporary viewpoint.

Chapter 3

Etiological Teleosemantics on intentionality

3.1 Teleosemantics

3.1.1 Teleosemantics in context

Let's start by first briefly locating teleosemantics within the history of the philosophy of language and mind (cf. McGinn, 2015; Soames, 2010, for an introduction). Questions concerning semantics and the nature of meaning make up the core of the analytic tradition in the philosophy of language. Teleosemantics occupies a place within this tradition. Specifically, since the beginnings of the contemporary philosophy of language, there is an ongoing and intense debate about internalist and externalist theories. As Millikan states, "naturalistic teleological theories are 'externalist' theories of mental content" (Millikan, 2003, 1). I will briefly introduce the internalist and externalist perspectives taking as a reference their original proponents, Gottlob Frege (1848–1925) and Bertrand Russell (1872–1970) respectively, in order to see why a teleological notion is needed for an externalist theory of content.

It is probably safe to say that it was Gottlob Frege who first put semantics at the top of the agenda. Although he was never interested in the nature of psychological states, but rather in the semantics of formal languages and the problems that arise in natural language as a consequence of not being such a perfect system as those of mathematics or logic, his proposals have nonetheless had an impact in mentalistic models. It is in this context where he introduced the distinction between sense and reference (Frege, 1892). I am not going to get into the details of Frege's theory, insofar as it, as noted, was not a theory about inner psychological states, but about the meaning of external words, and formal symbols. The cognitive dimension of semantics and its explanatory role in behavior only came

in the mid 20th century with the rise of the cognitive revolution. Frege's theory is still interesting, however, to the extent that it is the source of a number of spin-offs within the philosophy of cognitive science, that we could aptly label as Neo-Fregean, aimed at the understanding of contentful intentional states. In a nutshell, Frege's distinction between sense and reference was meant to explain how reference is determined. As he proposed, the sense of an expression determines its reference, and the sense of an expression was defined as its *mode of presentation*. In other words, the reference of an expression is presented in a particular way by those parts that constitute the expression. To use an example: I can refer to the current president of Uruguay at the time of writing this chapter (January 2022) in many ways: using the name 'Luis Lacalle Pou', the expression 'the 42nd president of the República Oriental del Uruguay', or by any other expression that has the current president of Uruguay as its reference. Such a way of presenting the reference is the meaning of the expression. As we can appreciate, different expressions with different meanings can have the same reference. In this vein, Frege explained how different expressions with different meanings can be co-extensional. What determines the content of an expression is the way it is presented —its meaning, or sense, as Frege called it. Crucially, to understand how the meaning of an expression is determined, he proposed the Principle of Compositionality: the meaning of an expression is the result of the composition of the meaning of the parts under a particular syntactic form. The meaning of *the president of Uruguay* is determined by the meaning of its constituents arranged under a particular syntactic form. The main moral is that the content of an expression is determined by the connection of the whole expression with other semantic/contentful units: in Frege's case, the components of the expression that —through compositionality— constitute the whole expression. The definition of content in terms of other contenftul units is the core of the internalist positions and it has its roots in the Fregean distinction between sense and reference.

For obvious anachronistic reasons, Frege's view is not rooted in cognitive science, but once Frege's ideas are transferred into a cognitivist framework, they entail that the connections between semantic units determine representational content. This is the idea that the intension —the inner semantic structure of representations— determines the extension —the reference to the world: Neo-Fregeans contend that intensions determine extensions (García-Carpintero & Macià, 2006). The connection between the parts of an expression —or the parts that constitute intentional states— determines what such expression is about. But Neo-Fregeanism may take many alternative forms. For example, some decades ago, a number of proposals maintained that the content of a representation is determined by the definition of the representation. The content of X is determined by other contentful representations that constitute the definition of X. Prototype

theories, by linking the semantic content of a representation with those typical properties associated with it, are also internalist, intensionalist accounts. Semantic networks, inferential/conceptual role semantics are also other cases where inner semantic connections constitute the buildings blocks of intentional states (cf. Fodor & Pylyshyn, 2015, for a critical introduction of Neo-Fregean theories).

As noted, teleosemantics is not intensional but extensional, so it abandons the view that defines content in terms of inner/mental relationships. The main reason for this is its commitment with naturalism. Simply put: one cannot explain content by appealing to other contentful stuff if one aims at naturalizing intentionality. This is connected with the distinction between derived and underived (cf. Section 1.2.2) intentionality. Underived intentionality explains content without invoking prior intentionality. A naturalist project must explain how intentional states could arise from non-intentional states and intensional theories appear to be incapable of doing that. Another way to put it is by following Fodor and Pylyshyn's critiques of the Neo-Fregean accounts (Fodor & Pylyshyn, 2015). According to them, if we explain content by appealing to other contentful stuff, we enter a vicious circle that cannot let naturalism get in: intentional states are explained by other intentional states.

But, how can we break this circle? Neo-Fregean theories do not seem to have the appropriate tools for doing it. Enter teleosemantics and, in general, externalist/referential theories of content. Teleosemantics offers a different view on representational content than those expounded by Frege and the Neo-Fregeans. Referential theories of content, principally in connection to natural language, are usually associated with Kripke's theory of proper names. In Kripke's theory (Kripke, 1980), contrary to Frege's proposal, the meaning of proper names is determined by a causal chain between the reference and the representation starting from the day of baptism. Content is determined by a connection between the name and the named. To be sure, Bertrand Russell's epistemological and semantic theory may also be seen as a referentialist manifesto of sorts. More generally, Russell's theory defends that all expressions must have a history that starts with environmental-mental relationships. Russell was a phenomenologist, so the basic semantic units for him were the so-called *sense data* (B. Russell, 1910, 2010). Thus, the central insight in Russell's theory that many other referential theories subscribe to —even though concerning proper names Russell is closer to Frege than to Kripke (B. Russell, 1905)— is that the building blocks of meaning and knowledge are environment-mind relations.

If we locate referential theories of content into a cognitivist framework (that is, if we are interested in mental representations and not just in expressions of natural languages), then some representations must be explained only by their non-intentional relations with the world. These theories put at the center stage

environment-mind relations. Therefore, If we can account for the content of some representations without intentional states being involved, then we can add a pinch of non-derived intentionality. If we add this pinch, the main step towards the naturalization of intentionality is done and we would be able to show how environment-mind relations could give rise to contentful representations without the involvement of intentionality.

Teleosemantics (like informational semantics; see below) was born in connection with referential theories of content, and it has usually preserved such a connection. Importantly, one can perfectly imagine how an intensional teleosemantics would look like. There is no necessary link between teleological theories of content and the internalist/externalist division. However, having said so, there is a clear reason why teleosemantics goes hand in hand with referential theories of content. This concerns misrepresentation. Intensional theories can solve it quite easily: misrepresentation occurs when the reference does not possess the properties that constitute the content of the tokened representation. As explained, the content is determined by other semantic units. Such semantic units refer to properties of the world. For instance, if the prototype of the representation DOG is to have four legs and be friendly, the representation DOG would be misrepresenting when it is tokened to refer to, let's say, a chicken. This is so because chickens do not have the properties that constitute the content of DOG. Intensions provide the criteria for a normative demarcation between proper representation and misrepresentation. As expected, the problem with this solution is that it is not suitable for a naturalistic program. Misrepresentations must be accounted for —at least as far as underived intentionality is concerned— without reference to other semantic stuff.

The need for a normative criterion based on natural teleology is connected with the abandonment of an internalist/intensionalist theory of content and the commitment with a referentialist account. If content is determined by reference, how is it possible to have representations without reference? How is the Causal Mismatch possible? Here, teleology enters the scene. Teleosemantics pursues an explanation of misrepresentation and the Causal Mismatch without appealing to prior intentionality: a teleological notion of function shall play this explanatory role.

To conclude this brief introductory section, let's note that I am not saying that Neo-Fregean theories are mistaken or inadequate; that depends on one's explanatory purposes. They are certainly inadequate to naturalize intentionality; at some point, you must preclude intensional explanations to tackle some representational, underived content. However, Neo-Fregean theories are central for explaining how it is possible, once the pinch of naturalized intentionality is added into the mind, to build complex, abstract, conceptual, and productive representational capacities. That much implies that a mix of internalism and externalism is necessary for

explaining mental representations. Without externalism we cannot anchor intentionality in natural science; without internalism, we cannot explain how complex knowledge and further representational capacities are achieved. In this thesis, I won't be concerned with the second issue. Instead, I will focus on understanding how environment-mind causal relationships can give rise to simple contentful representations.

3.1.2 The teleosemantics' core

In Chapter 1, I presented two connected riddles of the life sciences: Brentano's problem concerning intentionality and Kant's Puzzle concerning teleology. The core of any teleosemantic (naturalist) project is to solve Brentano's Problem on the basis of some solution to Kant's Puzzle. There are two central ideas in teleosemantics. First, intentional systems have teleological functions. Any cognitive function has a teleological dimension concerning the goal it is directed to fulfill. Second, intentional systems, like any other natural phenomena in the living world, fall under the scope of a natural theory of teleofunctions. Therefore, the analysis of teleofunctions in biology also encompasses that of cognitive teleofunctions. As Neander puts it:

> Teleological theories of mental content are intended to be naturalistic theories. They take seriously the idea that intentionality is a biological phenomenon. Further, their proponents think that, not only have cognitive systems evolved, but—and here is the crucial, controversial claim—in the fact of this evolution lies the solution to Brentano's problem. (Neander, 2008, 384-385)

As noted, the teleosemantics' core does not refer to any particular notion of teleofunctions. This is so because teleosemantics can take different forms depending on the position adopted concerning its two central ingredients: the *teleo* side and the *semantics* side. I will speak about the semantic side of teleosemantics and its variations in Section 3.3.2. As for the *teleo* side, different teleosemantic proposals arise depending on the theory of teleofunctions endorsed. Many theories of functions have been developed in biology in the last two centuries. Part of this project is about evaluating mainstream teleosemantics while simultaneously proposing an alternative path based on a different theory of teleofunctions.

Certainly, not all solutions to Kant's Puzzle are suitable for a solution to Brentano's Problem. There are two main constraints imposed by the teleosemantics' core. Firstly, we need a naturalistic theory of teleofunctions, so, for instance, Brentano's Problem would not be solved if it were based on Natural Theology. Secondly, the teleosemantics' core cannot be based on an eliminativist solution

to Kant's Puzzle. If we argue that teleological explanations are not necessary for biology and that Kant's Puzzle needs not be solved but eliminated, this would culminate with an unhappy end for intentionality: intentionality would not be naturalized but diluted.

3.2 Etiological Teleosemantics

3.2.1 Etiological teleosemantics and Selected Effect Functions

Mainstream teleosemantics is etiological and it is based on an etiological theory of functions.[1] The classical approach to etiology was developed by Larry Wright in his book *Teleological Explanations: An Etiological Analysis of Goals and Functions* (Wright, 1976); see also Ayala (1970) and Ruse (1973). An etiological theory of functions defines the function of a trait by looking at the effect it causes. There is an ample variety of etiological theories of function, but in mainstream teleosemantics, etiological functions are understood in terms of natural selection, what has come to be known as the Selected-Effect Theory of Functions (SETF), developed by Millikan (1984, 1989) and Neander (1991a, 1991b).

According to the SETF, the function of a trait is to do whatever it was selected to do during evolution by natural selection. This is clearly etiological. During evolution, the effects of traits in fitness values cause them to be selected; i.e. functions are explained by their effects. Moreover, this is also a *teleological* theory of functions. It posits that traits perform a certain function to fulfill a particular goal or purpose, according to the role that such a trait has had during its process of selection; this role explains why the trait is present in nature. Before getting into the details of the etiological solution to Brentano's Problem, I would like to make some preliminary observations in order to clarify the terminology, especially concerning the following question: are selected-effect functions (defined in terms of natural selection) the only possibility for etiological functions in biology?

My aim however is not to find an answer to this question but only to sort out some terminological issues for discussing mainstream teleosemantics. The question has to do with whether there could be etiological functions that are not based on natural selection. To be sure, in the literature etiological functions typically refer to the SETF defined in terms of natural selection. So, Etiological Teleosemantics

[1] Schulte (2020, 2273-2274) distinguishes between narrow and broad teleosemantics. Broad teleosemantics includes any attempt at explaining intentionality in terms of biological teleofunctions —i.e. any attempt to provide an account of intentionality based on whatever theory of natural functions. Narrow teleosemantics refers specifically to etiological teleosemantics, what I refer to here as mainstream teleosemantics.

(henceforth ET) is considered to be a teleosemantic theory based on evolutionary selected-effect functions. However, Shea (2018) defends, for example, that there could be etiological functions not grounded in natural selection. Accordingly, within this view, ET is not necessarily a matter of evolution. Shea accepts that there are etiological functions not only at the evolutionary level but also at the individual level (what grounds his *varitel semantics* —the variety of teleofunctions; cf. Section 3.3.1). For instance, he accepts etiological functions defined in terms of learning process and persistence. Somewhat surprisingly, however, when Shea posits persistence as a source of etiological teleofunctions, he cites Mark Bickhard as the proponent of this theory (Bickhard, 2000a, 2003). But, indeed, Bickhard is one of the principal critics of etiological functions. Shea does not seem to notice this terminological tension: Bickhard pictures his proposals as opposed to etiological functions yet Shea takes it to be an etiological account.

I shall solve this tension by taking a terminological decision. I will reserve the term *etiological teleosemantics* to refer only to those teleosemantic theories based on natural selection. As explained, this is mainstream teleosemantics, as it was developed since its inception. It is not the only possibility that scholars adopted, but it is still the classical way to do teleosemantics.

3.2.2 ET and the Causal Mismatch

Selected-effect functions have to do with those evolutionary processes that made certain traits being selected as a consequence of their adaptive advantages. ET includes cognitive functions under the view of the SETF. Cognitive teleofunctions are a product of evolution: a cognitive trait must do whatever it was selected for doing during natural evolution. Within this framework, the following question arises: what is then the solution to the Causal Mismatch offered by ET?

As noted before, teleology makes it possible the introduction of the normative ingredient in the cognitive sciences that is necessary in order to unknot the problem of intentionality. The central idea is, in a nutshell, that the Causal Mismatch could be accounted for insofar as cognitive functions are teleological functions. Cognitive systems are there to perform certain tasks oriented toward a specific goal. The causes provided by an intentional system to produce a particular behavior are related to the goal that such behavior is supposed to fulfill. This goal is defined according to the SETF.

With these considerations, ET can tackle the normativity of intentional states. Teleological functions define the goal of a trait, and the goal defines the normative criteria for determining whether the function is fulfilled or not by that trait. Therefore, and crucially, malfunctioning is possible and misrepresentation is allowed. Misrepresentation is the landmark of the Causal Mismatch, as explained before. If the function of a representational system is to represent whatever it was selected

to represent, then we can explain how misrepresentation is possible, that is when a representational system departs from its evolutionary function. Note that no prior intentionality seems be required in the explanation of the Causal Mismatch. This is the crucial advantage of teleosemantics when confronting Brentano's Problem with naturalistic aims. From this viewpoint, intentionality is rooted in a naturalist theory of teleological functions.

The core of this solution lies in the distinction between trait types and trait tokens. Selected-effect functions concern types, not tokens. Functions are defined at the populational level —the entities that evolve. Therefore, the normativity of content is established at the same level. Representational errors are possible to the extent that individual trait tokens are able to wrongly instantiate the type they belong to. If a certain token of a representational system —e.g. a particular frog's visual system— does not do what it was established by the trait type it belongs to (the evolved visual system of frogs), the possibility to misrepresent the world, make errors or produce maladaptive behaviors is explained. If the visual system of frogs was selected during evolution for representing flies in specifics environmental contexts, then a particular frog would be misrepresenting the world when it does not represent flies as flies when flies are flying out there, or when it represents flies when there are none. Representational errors are thus reduced to the mismatch between tokens and types.

Before diving into teleosemantics, I want to highlight an important issue stressed also by ET. It concerns the importance of the historical dimension of teleological functions —and, therefore, of normativity too. The SETF is usually considered to be the rival of the so-called Cummins-functions (Cummins, 1975). Cummins-functions refer to the causal role played by a trait in the current functioning of a system. Although I will have more to say about Cummins's insight later, there is an important point that already needs to made explicit here. Cummins-functions are defined at the individual level. They concern tokens and the role such tokens play in the system they are a part of in a particular moment. Cummins-functions do not have a historical dimension: they just require an analysis of how a system operates and how its parts interact. Cummins-functions have no teleological dimension —they are not teleofunctions: there are no goals involved in defining functions just by looking at the current interactions of a system. What etiology adds to the study of biological function that surpasses Cummins's analysis is, as Neander (1991b) stressed, a historical dimension. What history adds is a reference to the process of function-establishing. The idea is that *goals are defined by the effects of past events*. If such past events are adaptively biased —be that by natural selection or by individual development— the trait obtained —by evolution or development— would have a teleological dimension concerning those functions that contributed to the trait being adaptive.

This is at the core of the solution to Kant's Puzzle on teleology within the MS framework (cf. Section 2.3.3). Teleology explains by citing the effects of a current activity insofar as the effects were defined in the past. Here, the Causal Asymmetry Principle is respected. Teleology is rooted in a causal chain that comes from the past to the present, as modern science demands. That is why Shea states that "without the historical angle we would be back to the mystery of teleological causation, the mystery of how it is possible to explain a cause in terms of the type of effect it is likely to produce" (Shea, 2018, 59). I shall call this the *Historical Desideratum*:

(HD) **Historical *Desideratum***
Without a historical dimension concerning an adaptive bias operating in the past —be that in prior evolutionary or developmental stages— there is no teleological analysis in the present.

As explained by Macdonald and Papineau, the SETF can deal properly with the HD:

> On this account of function, functions are the upshot of prior processes of selection. A trait has a function if it has been designed by some process of selection to produce some effect. In the central cases, where the traits in question are biological *adaptations*, the selection process will be non-intentional natural selection. An effect of a trait counts as its function if the trait has a certain history: in the past possession of that trait produced the relevant effect, which in turn had the consequence facilitating the reproduction of items with that trait. In such cases, it is natural to adopt teleological terminology, and say that, in the normal case, the trait exists *because* of an effect the trait can produce, or *in order* to fulfil its function. (Macdonald & Papineau, 2006, 10–11; emphasis in the original)

3.2.3 Replication, externalism and causalism in etiology

This subsection aims to show that ET has its roots in the framework of the MS. In Section 2.2, I defined three central pillars of the MS as it was construed in the 20th century. I will argue that such pillars are also present in the SETF. The second part of this thesis will be devoted to analyzing the answers to Kant's Puzzle and Brentano's Problem presented in Chapter 2 and Chapter 3, respectively. This scrutiny will be approached by looking directly into the pillars of the MS. So it is central to clarify first the connections between the MS and ET. Our first target will be causalism.

As explained in Chapter 2, the idea of evolutionary processes as governed by populational forces was forged during the MS. It is clear then that ET rests on populational causation. The SETF systematically refers to Sober's insight on selection-for effects. Sober's proposal, as noted, is a classical locus of the causalist view on natural selection. Millikan, as most etiologists, appeals to Sober's distinction in order to define her notion of proper function: "[o]nly if an item or trait has been *selected* for reproduction, *as over against other traits*, *because* it sometimes has a certain effect does that effect count as a function" (Millikan, 1993, 35–36; emphasis in the original). In this sense, if the SETF defines functions in terms of the effects of a trait in evolutionary processes, such effects are not selection-of effects, but selection-for effects. Neander also makes this explicit: "[o]n an etiological theory, functions are what entities were selected *for*. Mere selection *of* a trait is not enough to confer a function on it" (Neander, 2017b, 132; emphasis in the original). Natural selection is the causal force that establishes etiological functions: "Selection does more than merely distribute genotypes and phenotypes[…]: *by* distributing existing genotypes and phenotypes it plays a crucial causal role in determining which new genotypes and phenotypes arise" (Neander, 1995a, 585; emphasis in the original); or as Artiga (2021, 53) recently stated, with an explicit reference to Sober: "for a trait T to be selected for F […] F must be an effect of T that causally contributed to success (Sober, 1984, 97-102)." The adoption of a causalist position by etiologists is quite expected. Selected-effect functions are defined as types. As explained, trait types —defined at the populational level— are the source of teleofunctions. If the process that attributes functions to such types is not causal, the causal basis of teleology would be missing.

The commitment to Explanatory Externalism is also present in ET. Under the MS, the environment, not the organism, is responsible for biasing a population in an adaptive way. The organism has no explanatory role in selection processes (cf. Section 2.3.2). As selected-effect functions are based on natural selection, Explanatory Externalism is also present in the analysis of evolutionary functions. The function of a trait is to do whatever it was selected to do by the environmental conditions that made it fitter than other variants during its evolutionary history.

A direct consequence of this commitment is that representational systems must be explained by natural selection. That is a mandatory requirement. If representational systems were not explained by natural selection, then the theory of selected functions would remain inert at the time of explaining the teleofunctions of cognition. Correspondingly, no other evolutionary adaptive forces may be relevant. That is, if other adaptive processes beyond selection took place during the evolution of cognition, then such processes cannot be accounted for by the SETF. So ET must avoid any other source of adaptive evolution beyond natural selection. Therefore, the explanatory force of ET crucially hinges on the adequacy of

understanding the evolution of cognition as guided by external selective processes.

Finally, replication is also present in the SETF. As Papineau has stressed, "[c]entral to the etiological account is the idea that individuals gain functional traits as a result of being *replicated*" (Macdonald & Papineau, 2006, 12; emphasis in the original). This is particularly clear, for instance, in Millikan's account of functions, as it was originally presented in her *Language, Thought and Other Biological Categories* (Millikan, 1984, Chs. 1 and 2). In her view, the notion of copy is central in defining proper functions. The idea of replication rests on the irrelevance of how the replica or copy is actually made —which developmental processes construct the new tokens. In etiological accounts, the attribution of functions is made on the basis of fitness values without taking into account how such traits are constructed. ET cares about the functioning of adult phenotypes insofar as they were the target of past selection. Selection plays its card generation after generation with replicas that have become biased towards adaptations. Vehicles are not a source of proper functions.

3.3 Within teleosemantics

In this section, I will get deep into teleosemantics to present different proposals, their points in common, and their disagreements. I will also introduce in detail the thought of its main characters: Ruth Millikan, Karen Neander, David Papineau, Nicholas Shea, Fred Dretske, and Mark Bickhard.[2] The presentation will follow the wake of two principal questions concerning the two core ingredients of teleosemantics:

(i) The teleological question, or where do teleological functions come from?

(ii) The semantic question, or what theory of content is posited?

On the one hand, question (i) asks about the different possible sources of teleofunctions. As noted in Section 1.1.2, there are three temporal dimensions in biology: the evolutionary dimension, the developmental dimension, and the physiological dimension. Mainstream teleosemantics —that of Ruth Millikan, Karen Neander, David Papineau, and Nicholas Shea— is rooted in the evolutionary level, whereas Fred Dretske classically defended the developmental time scale as the proper source of cognitive teleofunctions, and Mark Bickhard has argued for norms and functions at the physiological time scale. However, as it will be stressed later, to the extent that an account of functions exclusively based on evolutionary criteria

[2]There are other scholars who have made relevant contributions to teleosemantics such as Artiga (2013, 2014b, 2020); Artiga and Martínez (2016), Schulte (2015, 2018, 2020), Nanay (2014), and Martínez (2013a, 2013b).

is a controversial issue, most teleosemanticists tend to adopt a pluralist position concerning the sources of teleofunctions. Question (ii), on the other hand, refers to the classical distinction in teleosemantics between consumer-based teleosemantics —Ruth Millikan and David Papineau— and producer-based teleosemantics —Fred Dretske, Karen Neander, Nicholas Shea, and, up to certain extent, also Mark Bickhard. So let's now move to present these authors' ideas and their answers to the two aforementioned issues. My aim is descriptive, and the exposition of these teleosemantic projects is meant to set the scene in order to thereupon discuss which ideas need to be rethought, which ones need to be rejected, and which ones will be endorsed and further developed in this thesis. The granularity of my presentation is related to the scope of the analysis. If further details are needed, I shall introduce them as I proceed.

3.3.1 The teleological question: Pluralism about teleofunctions

ET sees in natural selection the primary source of teleofunctions. Ontogenetic functions are usually perceived as an extension of evolutionary functions. While they are not relegated to the status of lower class citizens, ontogenetic functions are considered to be a complement to evolutionary functions. This is so insofar as ontogenetic functions play the important role of solving certain complex issues that are problematic without incorporating an individual-level ingredient. There are principally three problems that are difficult to address exclusively from an evolutionary perspective:

(i) The problem of novel contents;

(ii) the context-dependence of most representational capacities; and

(iii) the variation in representational capacities between different groups with the same evolutionary history.

Point (i) refers to the implausibility of assuming evolutionary histories for novel representations, such as ELECTRON or WI-FI. Natural selection seems not to have had enough time to select for these novel representations, so another source for teleofunctions must be performing this task. Point (ii) concerns the obvious dependence on environmental inputs of representational development. This is the classical locus of some long-standing dichotomies —such as e.g. innate vs. learned— endorsed by most etiologists (cf. Section 5.2). If some representations are learned, then evolution cannot provide a full-fledged account of representational content. Finally, issue (iii) concerns the fact that there are clear differences in representational capacities among humans as a consequence of having developed

in different niches. Natural selection seems not to be capable of capturing these representational differences, therefore an ontogenetic ingredient must necessarily be present. Let's start by sketching Millikan's theory of functions first.

Ruth Millikan is the main character in teleosemantics, a framework she inaugurated in her celebrated *Language Thought and Other Biological Categories* (Millikan, 1984), and which she has continued to cultivate with a vast amount of work in the field through the years (Millikan, 1993, 2004, 2017). She was the first one to link an evolutionary account of proper functions with the analysis of representational systems. Indeed, the defense of proper functions at the evolutionary level as a consequence of natural selection and its connection with semantic content was originally introduced and systematized by Millikan. However, additionally, she also posits three other possible origins for teleofunctions beyond natural selection: derived functions, learned functions, and functions based on non-genetic modes of inheritance —particularly, cultural inheritance.

The first extension beyond selected-effect functions concerns her complex and layered view on functions. According to it, certain functions are derived from proper functions. Derived functions are thus not directly selected during evolution but are grounded in selected-effect functions. Secondly, already in her earliest account of teleosemantics, Millikan posited ontogenetic functions, principally based on learning processes, such as operational conditioning, practical reasoning, trial and error (cf. Millikan, 2000a, 86 and Millikan, 2006, 102, 103). Although learning is an ontogenetic process, learning mechanisms are evolutionary selected: "In the case of innate abilities, no matter what dispositions a mechanism happens to have, what determines its abilities is what it was selected for doing. In the case of learned abilities, what natural selection selected for was the ability to learn in a certain way. It selected for mechanisms that became tuned through interaction with the environment to do things of useful kinds" (Millikan, 2000b, 63). So systems of representation are the result of classical interactivism between ontogenetic and evolutionary resources: evolved proper functions interacting with environmental inputs during individual development. Finally, another extension of selected-effect functions defended by Millikan concerns extended —non-genetic— systems of inheritance (Millikan, 1984). Particularly, Millikan adopts Dawkins's memetic theory of cultural inheritance (Dawkins, 1976). So beyond the classical genetic transmission of proper functions, culture also operates as a source of proper functions by transmitting down through generations important information for cognitive development. Specifically, Millikan stresses the importance of cultural systems of inheritance to explain the semantic dimension of words in natural languages.

David Papineau is also an important figure in teleosemantics (Papineau, 1984, 1987, 1998). In many respects, his proposal is intimately connected with Millikan's.

Concerning the sources of teleofunctions, Papineau defends a classical etiological view based on natural selection, but, like Millikan, he also accepts other natural processes that introduce functions into cognitive systems: "I shall distinguish three ways in we can have etiological functionality in nongenetic traits [sic]. The first, emphasized by Ruth Millikan, appeals to a many layered account of functions. The second involves nongenetic selection in learning. The third depends on the intergenerational inheritance of nongenetic items. Together these three processes greatly *expand the range* of items that can possess etiological-selectional functions" (Papineau, 2017, 118; emphasis added).[3]

The position of Karen Neander is similar. Natural selection is the main source of teleofunctions (Neander, 1991a, 1995b), yet she also accepts an ontogenetic account for the teleofunctions of some representations (Neander, 2007, 2017b): "While the functions can be determined by phylogenetic natural selection, operating on a population over generations, they can also be refined or altered by ontogenetic processes involved in development or learning. In the case of ontogenetic processes, these might be selection processes, or they might be processes that are not selection processes but are, nevertheless, adaptations for further adapting the individual to its environment" (Neander, 2017b, 153).

Nicholas Shea's teleosemantics is set forth in his recent *Representation in Cognitive Science* (Shea, 2018). A strength of his theory is that he supports what he calls *varitel* teleosemantics, that is, a pluralistic stance both about the sources of teleofunctions and the determination of semantic content. His theory of teleofunctions —concerning representation in cognitive science— is a mixture of many proposals in teleosemantics. He identifies four sources for teleofunctions: deliberate design, learning, persistence, and natural selection. Deliberate design occurs when external agents set up a system to function in a certain way. These systems can be explained by the presence of derived teleofunctions. Persistence, an item also considered by Bickhard (see below), has to do with the capacity of self-preservation of individual systems. These processes, present at the (neuro)physiological level, are goal-directed towards the fulfillment of the system's needs in order to stay fit and alive. Learning processes attribute functions by establishing adaptive organism-environment relations tuned during development by the dynamic interaction between behavioral feedback and adjustments effected by the system itself, such as operational conditioning. Natural selection, as it should be clear by now, is the locus of selected-effect, etiological functions.

There is, however, another interesting ingredient in Shea's account. Elsewhere, he has developed a compelling and sophisticated etiological theory of evolutionary functions (Shea, 2007b, 2011, 2012a, 2013). The particularity of Shea's etiological

[3]In a recent paper with Justin Garson (Garson & Papineau, 2019), they appeal to ontogenetic functions in order to solve the problem of novel content.

theory has to do with two facts. First, he accepts and incorporates many ideas that fall beyond the standard evolutionary framework assumed by the MS, such as the parity thesis (Griffiths & Gray, 1994) and extended inheritance systems (Jablonka & Lamb, 2014), to which we turn presently. Second, he proposes a teleosemantic analysis well beyond cognition. His theory of inherited representation combines both teleofunctions and representational content with the aim of embracing the development and evolution of all living beings. I shall present and discuss in detail his view on inherited representation in Section 5.2.3.

Last but not least, we have the work of Fred Dretske and Mark Bickhard whose approach to teleosemantics may be safely tagged as non-classical. Some of their proposals will be adopted —with qualifications— in the third part of this thesis. The major work where Dretske articulates his teleosemantic framework is *Explaining Behavior* (Dretske, 1988; but see also Dretske, 2001, 2004). His pluralism on teleological functions is indeed central in his theory. While in other approaches different sources of teleofunctions are put forward to address different complex issues in teleosemantics, in Dretske's theory teleofunctions play the role of differentiating between kinds of representational systems; that is, his taxonomy of representational systems is based on his taxonomy of teleological functions.

Dretske identifies three types of representational systems. Type 1 includes those systems the activities of which are externally designed. Type 2 regards natural signs, present, for instance, in plants and some innate cognitive representational systems. Finally, in Type 3 systems, content enters the scene to posit representational capacities bearing semantic information. Dretske's pluralism about functions helps him at the time of differentiating these representational systems. As noted, Type 1 systems have extrinsic, derived proper functions, a view, for instance, also endorsed by Shea. As for Type 2 systems, he adopts a classical evolutionary approach, where natural selection is the source of the teleofunctions of natural signs. Dretske's idiosyncrasy becomes manifest with Type 3 representational systems, because the source of teleofunctions here is not evolution but individual learning —i.e. a developmental process. Representational content is established during individual processes of learning coupled with the environment and mediated by feedback processes acting in the course of development.

Finally, in his theory of mental content, Mark Bickhard develops a notion of teleofunction based on persistence (Bickhard, 2000a, 2003, 2009a, 2009b). Persistence has to do with the capacity of self-maintenance. This is clear at the cellular level. Cell metabolism makes it possible to preserve the unity of the cell by taking in energy and matter from the exterior and transforming them into metabolic resources for reproduction and maintenance. Persistence involves teleofunctions. Each element of the living system plays a specific function in the process of self-maintenance. Crucially, this function is directed towards the goal of preservation.

If a particular trait does not perform its proper function, degradation ensues and the goal would remain unfulfilled. As noted, this notion of teleofunction lies at the physiological level —it does not concern development or evolution.

3.3.2 The semantic question: producer- and consumer-based teleosemantics

There are two big families of teleosemantic theories: consumer-based approaches —represented by those of Millikan and Papineau— and producer-based accounts —defended by Neander, Shea, and Dretske (cf. also Stampe, 1977). The former explain the content of mental representations by looking into the effects of such representations —that is, how the system uses a representation to generate a certain behavior, regardless of how it is produced. Consumer-based accounts state that the content of a "representation depends on the output of the representation, on what behavior it prompts, and not on the input to it, on what circumstances cause it" (Papineau, 2017, 108). Alternatively, producer-based accounts focus on those processes that produce mental representations.

The separation between producers and consumers is central in Millikan's theory. The producer concerns the process that generates a particular representation, while the consumer refers to the use of the representation. Millikan's theory defines the content of a representation on the basis of its consumers. Accordingly, the use of a representation to produce a certain behavior determines the adaptiveness of such behavior. Therefore, the use of representational devices is connected with the selective advantages of a particular behavior. Consumers that have increased their fitness values during evolution tend to be selected. Even thought the correspondence between representation and reference is central, how such correspondence is achieved is not an explanatory aim of consumer-based teleosemantics: "a useful 'correspondence' between representation and represented does indeed occur when the biological system functions properly, but how this correspondence is brought about is not definitional of the representing relation" (Millikan, 2003, 3). As Neander suggests, consumer-based theories could be labeled *benefit-based* theories, insofar as they "link content to the benefit to the creatures (or to the consuming systems) that accrues from the use of a representation" (Neander & Schulte, 2021).

Millikan's theory is centered on animal communication as one of the main situations where representational systems are involved and a teleological analysis can be provided. From bee dancing and vervets' alarm calls to human language, the content of a representation is defined by how this representation is used to generate a particular behavior, such as communicative behavior: the content of alarm calls voiced by vervet monkeys means "eagle" because the calls make monkeys behave appropriately —i.e. adaptively— in the presence of eagles.

Papineau's theory is also aligned with a consumer-based framework. Yet, unlike Millikan, he restricts genuine contentful representations to those systems exhibiting desire/belief intentional states. Desires, in collaboration with beliefs, are responsible for producing intentional behavior. As he defines it, the content of desires resides in the effects that they produce —not in what produces them: "some past selection mechanism has favored that desire — or, more precisely, the ability to form that type of desire — in virtue of that desire producing that effect" (Papineau, 1993, 59). His emphasis on desires is due to their assumed causal role in producing behavior: the mechanisms underlying desire-states are the locus on which natural selection acts to select among different behavioral outputs. So there is a direct connection between the fitness consequences of behavior and the internal states —desires and beliefs— that are *used* by the system to produce a particular behavior.

The main source of producer-based teleosemantics is informational teleosemantics, originally represented by Dretske's Indicator Theory, and later developed by Neander (Neander, 2013, 2017b) and Shea (Shea, 2007a, 2018). At first sight, producer-based theories look radically opposed to etiological accounts of functions. As already pointed out, etiological theories concern the behavioral effects of a certain representation. Thus, ET should be about the effects of mental representations on the system, "the production of mental representations is irrelevant to their contents" (Neander & Schulte, 2021). Therefore, informational, or indicator theories stand "diametrically opposed" (Macdonald & Papineau, 2006, 7) to ET. Insofar as the main first steps in teleosemantics were covered by consumer-based theories,"[u]nmodified teleosemantics is entirely output-based" (Shea, 2007a, 409). But this is not entirely accurate. Teleosemantics can be —and it indeed was— refurbished. The central point is that functions must be selected functions, and selected functions are so because of their past contribution to survival and reproduction. Therefore, as Neander (2017b, Ch. 6) argued, there is nothing preventing the existence of selected *dispositions*; that is, the existence of adaptive functions designed to respond in a certain way. To define a function by its input does not block the possibility to explain the adaptive dimension of its output. Once the possibility of a producer-based account within an etiological framework was acknowledged, different accounts were developed on the basis that the content of a representation must be established in relation to those elements that trigger the process of producing a representation. This process includes the reference that triggers those perceptual and cognitive mechanisms responsible for building the corresponding —in relation to a semantic norm— representation.

Dretske's Indicator Theory (Dretske, 1981, 1988) is based on the correlation between the representation and its reference. An indicator bears information about those external things that caused it to be tokened. Representational con-

tent concerns the producers of representations/indicators —i.e. those aspects of the world that trigger the production of a representation. There is therefore a referential/correspondence relationship between the entity —the reference— and the representation/indicator. However, he correctly recognizes that many correspondence relationships in nature should not be treated as intentional. Following Grice (1991), he states that non-intentional correspondence relationships involve *natural sings*. For instance, smoke means fire in the sense that the presence of smoke corresponds with the presence of fire. Does fire represent smoke? A clear element seems to be absent here to achieve full intentionality: the possibility of error. If smoke is caused by fire, there could not be smoke without fire. But we know that one can have a representation without presupposing the existence of the representation. This is the locus of representational error. Natural signs do not allow the possibility of error, while intentional signs (representations) do have a normative dimension. As expected, the solution involves and appeal to teleology: "Dretske appeals to this fact to explain psychosemantic norms (the norms pertaining to mental content). His idea is that neural signals do not merely carry information, but are recruited or selected for doing so. In other words, they have information-carrying *functions*" (Neander, 2008, 394; emphasis in the original).

Neander (2006a, 2013, 2017b) and Shea (2007a, 2018) proceeded on Dretske's wake in order to argue for an informational teleosemantics. Importantly, informational teleosemantics is tied to informational talk in cognitive science. As explained in Chapter 1, information processing is explanatory central in the cognitive enterprise. This version of teleosemantics acknowledges the flow of information in cognitive systems, systematically posited by cognitive scientists, and adds a teleological dimension to such informational processing.

In informational teleosemantics, the Intentional Gap is typically presented as a clash between two different notions of information, which are related to the two kinds of sings (natural and intentional) discussed in Dretske's theory. Following Neander (2017b, Ch. 1), the first notion of information could be named *factive information*. It is based on the causal connection between two systems, in such a way that the presence of information crucially hinges on the existence of (at least) two systems. For instance, smoke informs that there is fire, dark clouds inform that there might be rain, and tree rings inform about tree age. In these cases, we have a causal connection between (at least) two phenomena (fire/smoke, dark clouds/rain; tree rings/tree age) in a way that one of them bears information about the other. As it can be noted, in some sense factive information involves aboutness: dark clouds inform about rain, or tree rings are about tree age. However, factive information does not attain the level of intentionality. Although some sort aboutness is present in these cases, what is absent here is the normative dimension. Therefore, misrepresentation (or misinformation) is not possible. And it is

quite clear why misrepresentation is not possible: if the presence of factive information depends on the existence of two causally interrelated phenomena, factive information would not be generated if one of these phenomena did not occur. As explained, misrepresentation takes place once there is a representation without a reference —i.e. when a representation bears information about something that is not there. That is why this notion of information is factive: it rests on the factual existence of a sign and its reference; factive information "is *factive* because nothing can [...] carry the information that some state of affairs, P, is the case, unless P is in fact the case" (Neander, 2017b, 7; emphasis in the original).

The situation is different for the other notion of information: semantic information. Here intentionality occurs. Like factive information, semantic information has aboutness: mental representations are about a certain state of affairs in the world. However, contrary to factive information, semantic information does have a normative dimension; i.e. there is room for misrepresentation: mental representations could perfectly be about states of affairs that do not really exist. If a representation or sign bears semantic information about its reference, this does not necessarily mean that some referent must exist. As expected, semantic information confronts us with the problem of the Causal Mismatch.

Factive information is based on causal relationships between different phenomena and it is usually modeled with Shannon's theory of information concerning the amount of information transmitted between two correlated systems that reduces the uncertainty of one of the systems, quite similar to Dretske's probabilistic account (Dretske, 1981). Unlike semantic information, factive information has relatively solid foundational bases. That is why factive information is ubiquitous in different scientific fields, from the thermodynamic perspectives in physics (e.g. Smolin, 2001) to biology (e.g. Crick, 1958) and neuroscience (Neander, 2017b). However, its robustness is a consequence of its incapability to account for misrepresentation and, accordingly, for the Causal Mismatch. In other words, factive information is not suitable for naturalizing intentionality.[4] Informational teleosemantics is about bridging the intentional gap and reaching semantic information by introducing a teleofunctional dimension to factive information: "the aboutness of content originates in the aboutness of information, and the norms of content originate in the norms of proper functioning" (Neander, 2017b, 125).

Before closing up, let's first present a classical and illustrative way to make explicit the difference between both families of teleosemantic frameworks, by introducing the well-known example of the frog. Frogs eat flies. To do so, they catch them with their tongues. This of course requires complex coordination between perceptual and motor capacities. The frog's case became a locus of debate be-

[4]But see Skyrms (2010) for an attempt to naturalize intentionality exclusively based on factive information.

cause it seems that in this dialogue —between what produces the representation (the perceived fly) and what consumes the representation (the snatching of the fly)— underlies a case of error. As stressed by different studies, frogs also try to catch things that are not flies but look like them, such as small and black dots moving on a screen. Enter etiology. A notion of natural norms for representational content is needed to explain why we are in the presence of a cognitive error. Consumer and producer accounts differ on which is the proper function of the representational systems. Importantly, different functions entail different contents. Consumer-based accounts would say that the content of the frog's representation is fly food, insofar as it is the property of being nutritive that made the representation advantageous during natural selection. To the contrary, aligned with producer-based accounts, Neander (2017b) argues that the content should be a small, black, moving thing, insofar as the representation is produced whenever such things occur. There is no consensus or unification, yet this is an internal discussion in teleosemantics, not an external dispute.

I believe, and shall defend in this thesis, that there are two interesting points in ET. Its referential semantics, particularly in its informational versions, and the endorsement of the teleosemantics' core. In Part III I will support both the teleosemantics' core and informational teleosemantics. I recognize that the semantic question was superficially treated here. This is so because Part II is about a specific problematic question in ET: the teleological question, i.e. etiology itself. My aim for the second part, specifically for Chapter 5, is to evaluate the etiological solution to Brenatano's Problem in connection with its biological foundations. So I will postpone a defense and inquiry into the semantic question while I propose certain challenges to the SETF.

3.4 Summary

In this chapter I introduced, situated, and described teleosemantics. First, I pictured the core of teleosemantics' explanatory strategy towards a solution to Brentano's Problem; thereafter I presented ET as a specific manifestation of such teleosemantics' core; finally I introduced some classical and important proposals in the history of teleosemantics.

I started by connecting teleosemantics with referentialist theories of content. Such a link is justified by two interconnected issues: (i) the inadequacy of intensionalist theories to naturalize content, and consequently, (ii) the need for a teleological notion of function in order to supply referentialist theories of content with a normative dimension in such a way that error may be explained without presupposing prior intentional stuff.

I continued by describing the teleosemantics' core. It concerns the central in-

gredients that any teleosemantic theory must have for it to be considered to be teleosemantic. The central idea is to appeal to a biological notion of teleofunctions in order to provide the grounds for dealing with the Causal Mismatch. As I pictured it, the teleosemantics' core ultimately amounts to assess a solution for Brentano's Problem based on a solution to Kant's Puzzle. In this sense, teleosemantics can take many forms depending on the solution to Kant's Puzzle that has been chosen. Indeed, this thesis will defend a teleosemantic project, based on an alternative, non-classical solution to Kant's Puzzle.

Even though there are different terminological uses, I referred to mainstream teleosemantics as ET: the idea that natural selection is the source of teleofunctions —the Selected-Effect Theory of Functions— and that this theory of teleofunctions applies to representational systems. I also showed how this view on biological functions has its basis in MS biology, or more specifically, how externalist-, replicator- and populational-based ideas are central to ET. Figure 3.1 on the next page illustrates the teleosemantics' core under an etiological view. The path starts with Kant's Puzzle and ends with the solution to Brentano's Problem. To solve Kant's Puzzle, it is needed to naturalize teleology. This is done by positing populational causes that do not involve any inversion of causal relations —i.e. without violating the Causal Asymmetry Principle. Populational causes define what means a trait type is being selected for. From this biological framework, the Selected-Effect theory functions grows and looks into the issue of intentionality to solve the Causal Mismatch. Once the Causal Mismatch is explained, intentional causation is finally naturalized, and consequently, Brentano's Problem is solved.

I concluded by presenting a number of classical proposals defended by different figures within teleosemantics. It was certainly just a rough outline and the reasons for this are twofold. First, I presented specific views within ET in order to evaluate their biological adequacy in future chapters. Here, my focus was on the *teleo* side of teleosemantics present in most teleosemantic projects. Second, the *semantic* side was only briefly presented for the very reason that I won't get into it till the last chapters of the thesis. Instead of criticizing teleosemantic accounts concerning the *semantic* side, I will embrace and develop different teleosemantic theories of content in Part III. There I will introduce in detail the theories involved.

In this part of the thesis I presented the two central and interrelated problems that I will deal with —Kant's Puzzle and Brentano's Problem, and the two classical and most accepted answers to them: The Modern Evolutionary Synthesis and Etiological Teleosemantics, respectively. Now it is time to evaluate these answers. Insofar as ET, and its naturalistic aims, is anchored in biology, specifically, in the MS, the discussion in the next part will be hosted principally in the (theoretical/philosophical) biological arena. Let's move on now to examine the adequacy of the picture drawn by the MS about the nature of living beings.

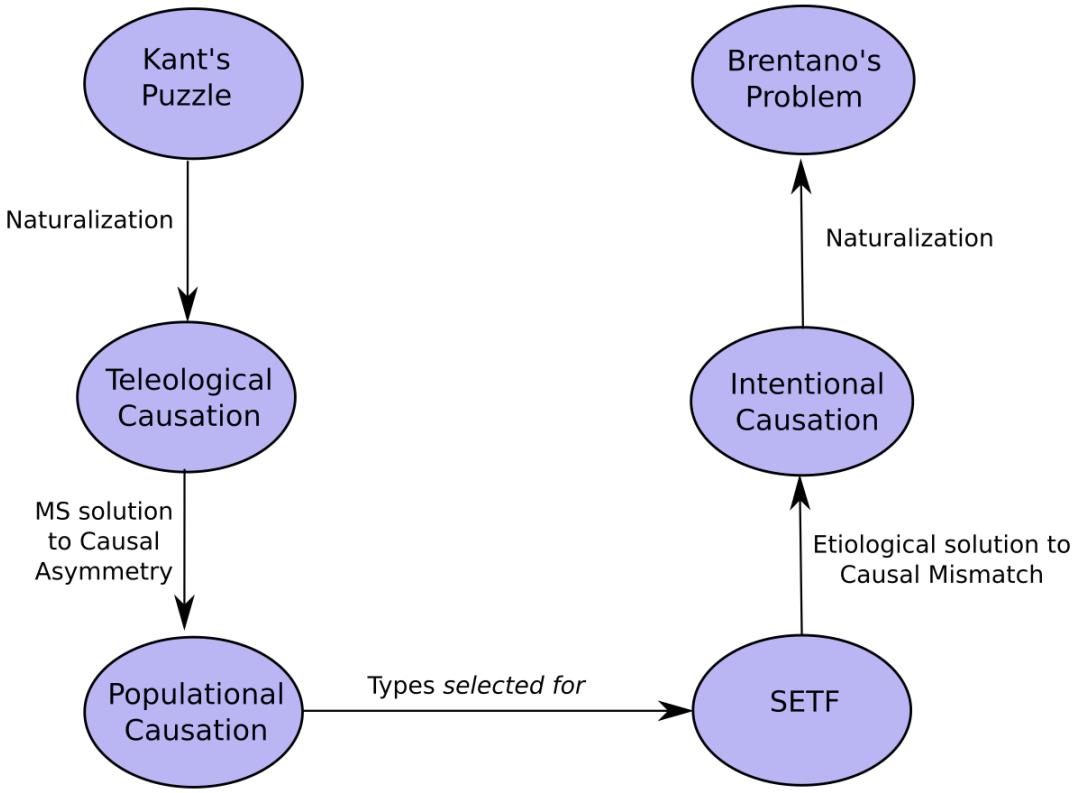

Figure 3.1: The teleosemantics' core: etiological solution.

Part II

Evaluating the answers

Chapter 4

Does the MS solve Kant's Puzzle?

> To take development seriously is to take development as our primary *explanandum*, to resist the substitution of genetic metaphors for developmental mechanisms… There is indeed good reason to believe that genetics reduces to development, and not the other way around.
>
> Jason Scott Robert 2004, 22

> The picture that emerges from recent developmental biology is that the stability and the mutability of organisms that are pre-requisites for adaptive evolution are consequences of the distinctive capacities of organisms, particularly as they are manifested in their development.
>
> Denis Walsh 2006a, 438

The purpose of this chapter is twofold. First, to present different challenges to the Ms in the last 40 years. These challenges are connected to the MS's pillars presented in Section 2.2. Each section of this chapter relates to one of the MS's pillars. Section 4.1, about the return of the organism in biological theory, is connected with Explanatory Externalism (Section 2.2.1); Section 4.2, focuses on the developmental view of inheritance and challenges Replicator Biology (Section 2.2.2); and finally, Section 4.3 discusses the idea of populational forces (Section 2.2.3) from a statisticalist reading of natural selection. Each section of this chapter could be read in connection with one of the three ingredients of natural selection: variation, inheritance, and fitness, respectively. To a certain extent, this is makes sense, since Explanatory Internalism is central to a theory of phenotypic variation; rejecting replicator ideas is central for moving forward towards a developmental theory of inheritance; and the statisticalist reading of natural selection is based on a specific account of fitness. But at the same time, all three ingredients are mixed and present in each part of this chapter. The reason is quite simple: the view of biology defended in this chapter is not a fragmented one (cf. Section 2.3.1).

All ingredients converge in development. So, inevitably, fitness, inheritance, and variation are intertwined, ontologically and epistemologically.

The second goal shall become explicit as I present the different challenges. It has to do with the sketch of a biological theory with a different approach than the one supported by the MS. Certainly, there is no solid and complete theory to contrast the MS, yet. The critical views against the MS constitute a cluster of ideas and experimental advances that fall outside the MS: what unifies them is indeed their rejection of some —or most— of the MS's ideas. So in this episode of the philosophy of science we find ourselves in *media res*. As Laland, Odling-Smee, Hoppitt, and Uller (2013, 807) said, "it is probably fair to say that these various lobbies currently more resemble a disorganized protest movement than a viable alternative government". Even though there are many alternative theories to the MS, such as the Extended Evolutionary Synthesis (EES) (Huneman & Walsh, 2017; Pigliucci & Müller, 2010), there is not a unanimous consensus as to whether such theories really amount to the conceptual revolution in theoretical biology they claim to be or whether the EES and other proposals are strong, coherent, and complete biological theories.

Moreover, as a consequence of this, there is still no consensus about what position theoretical biology should take with respect to the MS: should it be minimally modified, should it be extended or should it be replaced entirely? I am not going to discuss these questions here. I will just be introducing important ideas and experimental advances in the last decades that seem to go against the MS view of living beings, and reinforcing the idea that "a new biology [is needed] for a new Century" (Woese, 2004), because in the framework of the MS "an immense amount of biology was missing" (Lewontin, 2010). Having said so, it is relevant to note that intense disputes in theoretical biology have been taking place in relation to the validity and relevance of the challenges posed to the MS (cf. Futuyma, 2017; Gupta, Prasad, Dey, Joshi, & Vidya, 2017; Laland et al., 2014; Müller, 2017; Pigliucci, 2007, for different positions on this issue). Surely many of the challenges of the against the MS are not accepted —or even known— by orthodox biologists. Biology has not abandoned its MS niche; biological research runs in parallel to the philosophical issues discussed here. So maybe, there could be a revolution, but certainly, there is none, yet.

As there is no unified and alternative theory yet, I will opt to refer to these clusters of proposals, ideas, and research areas with the phrase *developmental turn* (henceforth, DT),[1] insofar as, in all cases, they deal with phenomena taking place at the individual level, and, as it will be clearer throughout the pages of this thesis,

[1] Jablonka and Lamb (2020) use the term *epigenetic turn*, although it might be adequate, I see 'developmental turn' as a broader label, insofar as it is not committed with some specific reading of what epigenetic means (cf. Griffiths & Stotz, 2013 and Baedke, 2018a, for discussions on the multiple meanings of epigenetics).

they adopt a 'developmental-first view': "Phylogeny is the derivational history of developmental systems" (Oyama, 2000b, 179). Even though not all areas deal with development *per se*, they all contribute to the understanding of the intrinsic capacities of living beings. As a tentative list, I suggest the following research areas as constituting the core of the DT collated, in no specific order, in Table 4.1 on the following page.

4.1 Explanatory Internalism: The explanatory role of organisms

This section challenges Explanatory Externalism (cf. Section 2.2.1); i.e. the idea that the unique adaptive force in evolution is natural selection. The reason for denying this MS's pillar is that internal organismal phenomena are also explanatory central to understanding evolution: "The black box [of development] is now being opened to provide a more complete picture of what really happens" (Bateson & Gluckman, 2011, 17). While in the MS, " 'selfish genes' in 'gene pools' are taken to be more important than organisms" (Reid, 2007, 11), now organisms are back in biological theory (Baedke, 2018b; Bateson, 2005; Huneman, 2010; Nicholson, 2014). As Waddington (1959, 1636) remarked, "[n]atural selection is very far from being as external a force as the conventional picture might lead one at first sight to believe". Internal forces are indispensable to account for the heritable variation in fitness. Specifically, the origin of phenotypic variations —a missing point in Darwin's theory, later filled in with a view based on randomness— needs ontogenetic explanations. This section has three parts. In the first one (Section 4.1.1), I introduce the post-genomic era and the new conception of genes in development that has come to the fore in the last twenty years. The second one (Section 4.1.2) stresses the many ways in which organisms adaptively regulate their ontogenetic trajectory by being sensitive to their external environmental conditions and their inner and complex dynamics. The last one (Section 4.1.3) is more theoretical and concerns the agentive and active role of organisms in evolution, motivated by the work of one of the principal architects of the DT, Richard Lewontin.

4.1.1 The post-genomic era

Probably the cardinal issue in the DT is the new conception of genes: their functions, definition, and role in development and evolution. I see this as the main motivation for rethinking the MS. Genes were put at the center of biological theory: they constituted the only source of heritable variation, and by being seen as encoding developmental programs, fitness differences obtained by phenotypic differences would be traced back to those genetic differences that produce traits.

Research Area	References
Evolutionary Developmental Biology (Evo-Devo)	Gould (2002); Love (2015); Minelli (2009); Minelli and Fusco (2008); Nuño de la Rosa and Müller (2021); G. P. Wagner (2014)
Ecological Developmental Biology (Eco-Devo)	Gilbert (2001); Gilbert and Epel (2015); Lewontin (2000); Sultan (2015); West-Eberhard (2003)
Developmental Systems Theory (DST)	Griffiths and Gray (1994); Johnston (2010); Johnston and Edwards (2002); Oyama (2000b); Oyama, Gray, and Griffiths (2001)
Developmental Psychobiology	Gottlieb (1997); Kuo (1976); Lehrman (1970); Michel and Moore (1995)
Embryology	Amundson (2005); Laubichler and Maienschein (2007); Robert (2004)
Cybernetics	Ashby (1991); Maturana and Varela (1980)
Molecular Epigenetics	Griffiths and Stotz (2013); Keller (2002); Moss (2003); Rheinberger and Müller-Wille (2018); Sarkar (2005)
Complex and Self-organized Systems Theory	Camazine et al. (2003); Goodwin (1994); Kauffman (1993, 2000); Müller and Newman (2003)
Systems Biology	Boogerd (2007); Kitano (2001); Noble (2016)
Extended Inheritance Systems Theory	Avital and Jablonka (2000); Jablonka and Lamb (2014, 2020)
Niche Construction Theory	Lewontin (1983b); Odling-Smee, Laland, and Feldman (2003); Scott-Phillips, Laland, Shuker, Dickins, and West (2014); Sultan (2015); West and King (1987)
Autonomous Systems Theory	Barandiaran and Moreno (2008); Bickhard (2000b); Kauffman (2000); Moreno and Mossio (2015)
Biosemiotics	Emmeche and Kull (2011); Favareau (2010); Hoffmeyer (2008a)

Table 4.1: A (not necessarily exhaustive) list of research areas associated to the Developmental Turn (DT), with references to some relevant literature.

In this subsection, I will challenge the idea of genes as encoders of developmental programs. In the next section, I will challenge the view of genes as the only source of heritable variation.

So, what ideas from the MS's view on genes should be revised? And, concomitantly, how genes should be understood? Let's start by separating two kinds of questions, one material and the other conceptual. I propose to formulate them as follows, where, below each question, I attach a hint of the answer I would like to develop:

1. **Material question:** Do genes really function according to the definition of 'molecular gene' advanced by Watson and Crick?
    - From the Molecular Gene to the Reactive Genome.
2. **Conceptual question:** Do genes still preserve their explanatory status?
    - From genes as the units of development to organisms a the units of development.

I call question 1 *material* because it concerns the material basis of genes. The question is therefore about the adequacy of the Central Dogma and the molecular gene as a theory about the material and functional constitution of genes. Do genes really function according to the definition of 'molecular gene' advanced by Watson and Crick? Can we characterize genetic activity as the Molecular Dogma states? I will argue that the notion of 'molecular gene' should be replaced by the notion of 'reactive genome'.

The conceptual question, 2, has different aims. This question concerns the fact that the explanatory role of genes in the MS did not depend on the material basis of genes. The transition from Mendel's view to Crick and Watson's discovery opened the door to significant experimental advances, but the explanatory logic behind genes as inherited units remained almost the same. This shows that the conceptual and material questions are relatively independent, to the extent that "the material form of the gene is inessential", whatever physical element it turns out to be, "the gene itself is a unit of information" (Griffiths & Stotz, 2013, 144). As the history of biology shows, changing the answer to the material question did not entail changing the answer to the conceptual one. The distinction between these two questions is a direct consequence of the fact that the centrality of genes in evolutionary biology was not tied to what genes materially are. So if there is any relevant and foundational conceptual change, it must be about the logic underlying abstract gene-talk. I will argue that the view of the gene as the unit of developmental control should be replaced by a view of the organism as the unit of developmental control.

The material question: reactive genomes

I shall start with the material question. An illustrative way to present the shift in genetics is by taking a look at the Human Genome Project (HGP). Allegedly, knowing a complete genetic sequence would give us the key to understanding an organisms' development and evolution. It would imply, for example, having access to the developmental program of the chicken that, as François Jacob claimed, is coded in the ovum; or as Rosenberg (2006, 61–62) recently defended, "the genes literally program the construction of the Drosophila embryo in the way the software in a robot programs the welding of the chassis of an automobile" (cf. also Rosenberg, 1997; Wolpert, 1994). Once the first human genome was sequenced at the beginning of the 21st century, the results were quite contrary to the expectations:

> Even later, advances in molecular biology, and propaganda for the human genome project, have allowed the mistaken belief that there must be a gene for everything, and once the genes and their protein products have been identified that's all we need to know. Instead, the completion of the genome project has clearly informed us that knowing the genes in their entirety tells us little about evolution. (Reid, 2007, 11)

Genes were far more complex than they were supposed to be. The aim of uncover the *Book of Life* (Lewontin, 2000) was downgraded once it became clear that it was nowhere to be found in the genes. Such a situation is not necessarily an undesirable one. To the contrary, it opened the possibility of studying an entity in all its newly discovered complexity. In this sense, "[t]he major theoretical achievement of the genome project was the refutation of its greatest expectation—that a mapping of the DNA base sequence would also be a map of all the interesting characteristics of the organism" (Lewontin & Levins, 2007, 82).

As pointed out by Griffiths and Stotz (Griffiths & Stotz, 2006, 2013; Stotz, 2006b, 2008; Stotz, Adam Bostanci, & Griffiths, 2006), the Human Genome Project represents a transition from the era of the molecular gene to the *post-genomic era* (cf. also DiFrisco & Jaeger, 2020; El-Hani, 2007; Keller, 2014; Moss, 2003; Perbal, 2015; Pigliucci, 2010; Rheinberger & Müller-Wille, 2018; Richardson & Stevens, 2015; Sarkar, 2005, 2006; Thorner, Hunter, Cantley, & Sever, 2014). Rather than being pictured as the units of developmental specificity, genes are now considered as part of a complex developmental matrix that interacts to produce traits. Developmental specificity —that is, those resources that contribute to the determination of a certain developmental outcome— is distributed among different levels of organization, not just the genetic one. On the one hand, yes, genes reside in genomes, and HGP unveiled the complex apparatus needed for protein expression. More specifically, it discovered that genetic activity can be influenced, activated,

or deactivated by the whole genome. The bi-directional flow of information and the multiple causal relationships genes participate in are not limited to the genome but scale up at different levels. But, on the other hand, crucially genes become active and functional in the context of the cell. This is why Barbara McClintock, in the speech she delivered on the reception of the 1983 Noble Prize for Physiology and Medicine, stated that the gene should be pictured as "a highly sensitive organ of the cell" (quoted in Keller, 2002, 33); see also Keller (2014) and Sultan (2015). The context-sensitiveness of genetic activity extends beyond cells, towards intercellular connections and even endogenous causes.

The Central Dogma pictured the molecular gene as carrying the whole information needed for development. The unfolding of this information is unidirectional, in such a way that strings of DNA sequence map into protein products. But all these ideas are in need of revision. Firstly, because the information needed for development cannot reduced to the genetic level, and other, genomic, cellular, extracellular, and exogenous resources are also central in development. Secondly, because DNA expression is not unidirectional or controlled by the DNA; rather, the activity of genes is dependent on their context and the expression of DNA may be altered by bidirectional pathways. Finally, because the mapping between DNA sequence and protein products is not so straightforward as the Central Dogma pictures it. Different DNA sequences, even from different parts of the genome, can act together in the development of the cell. The conclusion to the material question is that post-genomics has come to replace the molecular view of the gene promoted by the Central Dogma. Keller eloquently illustrates the range of problems behind the idea of the molecular gene:

> What is the causal role of a gene in the absence of environment? None is clearly the answer. Absent environmental factors, genes have no more power to shape the development of an individual than do environmental factors in the absence of genes [...] What we think of as its [DNA's] causal powers are in fact provided by the cellular complex in which it finds itself. It is this complex that is responsible for both the code that enables a sequence of nucleotides to be translated into a sequence of amino acids, for the replication of DNA, and for the intergenerational fidelity of replication; it is the cellular complex that makes possible all the chemical reactions on which these processes depend. By themselves, the entities we call genes do not act; they do not have agency. Strictly speaking, the very notion of a gene as an autonomous element, as an entity that exists in its own right, is a fiction. (Keller, 2010, 6)

It is not quite clear what genes are supposed to be in the post-genomic era, but the label that is often used to replace that of 'molecular gene' is *reactive genome*

(Gilbert & Sarkar, 2000; Keller, 2014). The notion of 'gene', as Keller (2010) stresses, is quite difficult to define once the action of DNA is functionally and structurally distributed throughout the genome. Many, along with Keller, have proposed that the very notion of a gene should be abandoned and that we speak about amino acid sequences instead. Be that as it may, the post-genomic gene represents a deep break with the molecular gene presented in the context of the Central Dogma.

The conceptual question: the organism as the unit of development

Do genes still have the same explanatory status in biology? This question concerns the possibility that even if the answer to the material question entails a rejection of the traditional notion of molecular gene, maybe the reactive genome still can be taken as the unit of development. I contend that the post-genomic era also calls for a re-conceptualization of the role of genes in biology. The main point has to do with taking the organism —from unicellular to multi-cellular organisms— as the proper unit of development.

What does it mean that the organism is the proper unit of development? The main idea is that the control of development —i.e. what determines the developmental path towards phenotypic outcomes— is the developing system. This entails that genetic activity is always tied to the needs of the whole developing organism. Several theorists within Developmental Systems Theory (DST), Eco-Devo, Developmental Psychobiology and Systems Biology support this view. Denis Noble, for example, asserts that "[i]n fact, the DNA just sits there, and occasionally the cell reads off from it a sequence that it needs, to get some protein produced" (Noble, 2006, 7). The moral is, therefore, that cells are not controlled by genes, but that cells have the custody of genes: "[t]he ghost [i.e. genes] in the cellular machine doesn't make the machine, and it doesn't make the machine run. The cell exists, and it runs 'by itself' " (Oyama, 2000b, 156). Or as Griffiths and Stotz (2006, 509) put it, "[g]enes are 'things an organism can do with its genome': they are ways in which cells utilize available template resources to create biomolecules that are needed in a specific place at a specific time". The holistic and complex nature of developmental systems appears to call for a *democratical* (Oyama, 2000a) distribution of explanatory roles among many developmental resources, rather than award their exclusive rights to genes.

The idea of the organism as the unit of development is connected to two issues that will be presented here but developed in detail in Part III: the notion of agency and the notion of the ontogeny of information. Agency concerns the capacity of living beings —from single cells to multi-cellular organisms— to regulate their own activity for adaptive reasons. By taking the cell as a minimal agent (cf. Section 6.3.2), the agentive capacities of cells connect with the idea of the cell

as the director of development. This view stands in contradistiction to the one promoted, for example, by Richard Dawkins Dawkins (1976, 1982). In Dawkins's view, genes are the agents and organisms the vehicles. But, genetic activity cannot be understood without a cellular context; the causal role of genes is dependent on such a context, and in such a radical way that "the bare genes in isolation are among the most impotent and useless materials imaginable" (West-Eberhard, 2003, 93).

The idea of the ontogeny of information, introduced by Susan Oyama in the early 1980s (Oyama, 2000b), puts the emphasis on the contention that the specification of developmental outcomes does not precede developmental processes. Ever since Weismann, the idea of the gene as the unit of development posits that information is "present, but unexpressed, in the constituents", and consequently, "the epigenetic building of a structure is not a *creation*; it is a *revelation*" (Monod, 1971, 7; emphasis in the original). In opposition to this, the idea of the ontogeny of information states that the specification of outcomes occurs during developmental processes as the result of the interaction among many developmental resources. Such a complex network of developmental resources is integrated and regulated by the organism itself according to the environmental context of development. Against Monod's quote, we must think of "development as creation, as *in-formation*" (Oyama, 2000b, 159; emphasis in the original).

The debate about whether information precedes development or whether information has an ontogeny is comparable to Schrödinger's views concerning the two alternative ways to understand the origin of order: the order-from-order strategy and the order-from-disorder strategy (cf. Section 2.2). Table 4.2 summarizes the different ways to understand order in biology in connection with the debate about the units of developmental analysis. Those who support that genes (whatever their material basis happens to be) provide the information for development tend to support an order-from-order strategy: the order present in living beings comes from an already ordered entity. However, once we remove genes as the source of order in development, we need to take the whole developmental system as the creator of order during development. The ontogeny of information, therefore, presumes that order comes from disorder. That the development of complex and adaptive traits comes from disorder.

Modern Synthesis	Developmental Turn
Order-from-order	Order-from-disorder
Genes as the units of development	The organism as the unit of development

Table 4.2: Alternative ways of explaining order: part I.

The post-genomic gene entails *The Death of the [molecular] Gene* (Gray, 1992), not as the rejection of the centrality of DNA in the living world, but as the re-conceptualization of its place in key biological processes, such as heredity, development, and evolution. As we will see presently, phenotypic variation cannot be reduced to the genetic level (Section 4.1.2), inherited material is not depleted at the genetic level (Section 4.2.1) and fitness differences cannot be traced back to genetic differences (Section 4.3.1). This new place of genes has a number of interesting consequences for biological theory. One of them is the issue I would like to turn to next which has to do with the reasons why the post-genomic era opens the possibility of understanding phenotypic variation beyond the gene-centric picture.

4.1.2 Development in context

As explained, the demise of gene-centrism opened the door for two challenges, one conceptual and other empirical:

> The conceptual change is a shift to thinking about the genotype as a *repertoire of environmentally contingent possibilities* rather than a single determined outcome. The practical innovation is to bring into experimental design the environmental variability that has been intentionally excluded from studies of both development and genetic variation under a strictly gene-based model of phenotypic determination. (Sultan, 2015, 20; emphasis in the original)

Sultan claims that we need to rethink development. What we are facing is therefore a conceptual issue. Once development loses its director (Griffiths & Knight, 1998), we need a new way to explain how the orchestra manages to play such beautiful music (Noble, 2006). As a result, the possibility of navigating into the complexity of development is opened up. That is a empirical issue. If developmental information is not encoded in genes in such a way that ontogeny is just about the unfolding of heredity, then different developmental resources must be interacting at different levels of analysis to produce phenotypes. In this subsection, I will present three central and interconnected research areas that came to the fore once the gene was removed from its pedestal: phenotypic plasticity, niche construction, and self-organization. I will show why these are central biological phenomena to understand the variation and adaptive origin of phenotypes.

However, before moving on, one should be aware of the fact that the phenomena introduced here —plasticity, niche construction and self-organization— are a source of intense disputes in contemporary theoretical biology. It is therefore important to recognize that an alternative, non-revolutionary interpretation of these phenomena is possible.

Phenotypic plasticity

Plasticity is "a ubiquitous, and probably primal phenomenon of life" (A. Wagner, 2013, 216). The importance of phenotypic plasticity in evolution results from the loss of developmental control by the molecular gene. Traits vary depending on the different environmental contexts in which development takes place. Phenotypic plasticity refers to the capacity of organisms to adjust and change their phenotypic traits according to their living conditions, in addition to finding alternative developmental pathways to preserve developmental outcomes (Bateson & Gluckman, 2011; DeWitt & Scheiner, 2004; Moczek, 2009; Moczek et al., 2011; Nijhout, 2003; Price, Qvarnström, & Irwin, 2003; Uller, Feiner, Radersma, Jackson, & Rago, 2020; West-Eberhard, 2005; Wund, 2012). As West-Eberhard (2003, 33) defines it, plasticity consists of "the ability of an organism to react to an environmental input with a change in form, state, movement, or rate of activity". Importantly, note that plasticity is taken to be a property of *organisms*. This is how plasticity is pictured by the DT, but certainly, as we will see later on, other proposals within the MS argue for treating plasticity as a property of genes.

Plasticity is associated with Norms of Reaction (NoR). The NoR of an organism with the same genotype shows how different phenotypic outcomes may arise in different environmental contexts. In Figure 4.1 on the following page are pictured different Norms of Reaction. Each plot represents the different possible phenotypic outcomes (P) that two organisms possessing particular genotypes (G1 and g2) can produce within different environmental (E) contexts. GxE interactions, where G is fixed and E is variable, give rise to a repertoire of possible phenotypic outcomes. Of course, not all GxE interactions produce different plastic traits (Schlichting & Pigliucci, 1998). A NoR may be robust if, in different environments, the same outcome is obtained, whereas a NoR would be plastic if the different environmental conditions in which development takes place elicit different phenotypic outcomes.

It is pertinent to emphasize the organismal dimension of phenotypic plasticity. As we will see later on, many MS biologists argued that plasticity should be understood as a property of genetic pools undergoing selection processes. Against this view, the developmental perspective of plasticity maintains that plasticity is a property of developing systems and that it cannot be explained just by looking at the level of genes; as Bateson and Gluckman (2011, 43) emphasize throughout their book, "the central elements underlying many forms of plasticity are epigenetic processes", with the net result being that "plasticity is an intrinsic property of organisms" (Sultan, 2021, 6). That much entails that plastic phenotypic responses during development are possible due to the capacity of the whole organism to adjust its developmental trajectory according to the different developmental contexts. In many cases, such changes require a complex interconnection among different parts of the organisms in addition to a sensorimotor coupling with the environ-

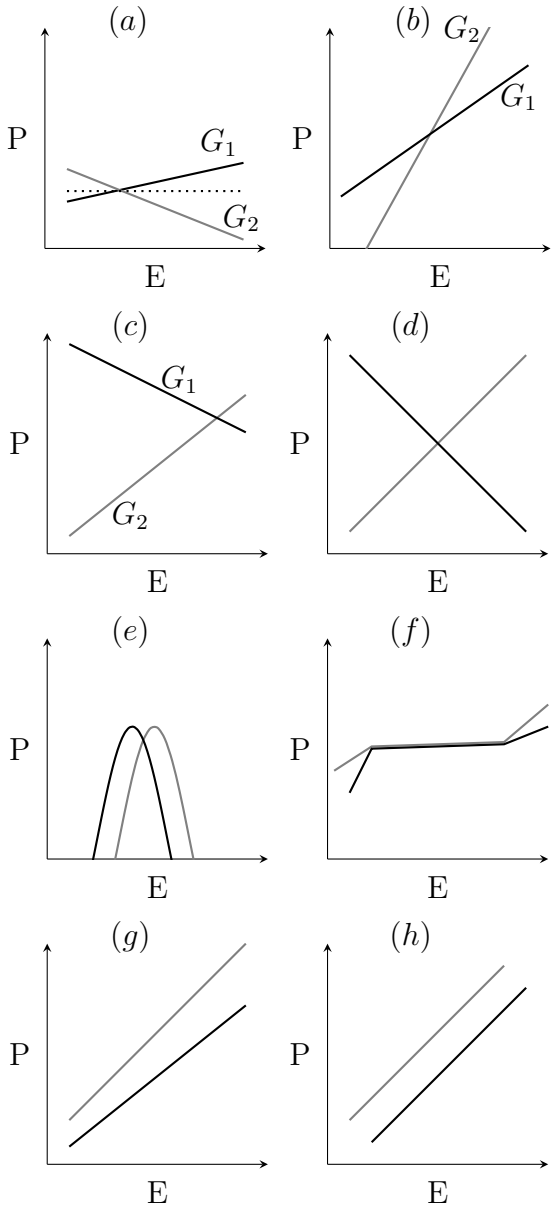

Figure 4.1: Some examples of different Norms of Reaction, where an ideal robust NoR is represented by the horizontal dotted line appearing in plot (a). For discussion of the significance of the plots shown here see Lewontin (1974a, 404–409); adapted from Lewontin (1974a, Fig. 1, 405).

ment in order to produce the variation suitable for acting in such environmental context, as witnessed, for example, in the renowned case of the Two-Legged Goat Effect discussed by West-Eberhard (2003, 51–54, 297–302). Plastic traits could be taken as organismal responses once we abandon the idea of a genetic program encrypting phenotypic outcomes. Variation is not prescribed in the genes, but it is brought about during the process of development.

Phenotypic plasticity is thus a central process for the introduction of trait differences —the rough material for selection. Therefore, plasticity deserves to be awarded a central role in evolutionary theory. Its centrality concerns the explanation of adaptive variations. Thanks to the context-sensitive and holistic character of development, phenotypic variation arising through plastic developmental processes is adaptively directed. To use an expression coined by Reid (2007), phenotypic variations (that is, the introduction of new or modified forms and functions in nature) emerge from the complex interplay between the many resources of development and physiology as "natural experiments" pursuing adaptive states.

A sharp contrast is manifested here with Explanatory Externalism. In the MS framework, natural selection is considered to be the sole adaptive force. In part, this is so because other processes were assumed to be blind and, particularly, the origin of phenotypic variation was considered to be the product of random processes, such that forces bringing about variation —drift and mutation— were never seen as adaptive forces. However, from the perspective discussed here, the situation is different: to the extent that phenotypic variation is taken to be an adaptively directed process obeying inner causes, Explanatory Externalism needs to be reconsidered.

West-Eberhard (2003) proposed a theory of evolution guided by variations arising through phenotypic plasticity which would then be stabilized in such a way that they would later be preserved and distributed by natural selection. Her view, known as *phenotypic accommodation*, was motivated by important (and neglected) proposals put forth in the late 19$^{\text{th}}$ and early 20$^{\text{th}}$ centuries, such as the so-called Baldwin Effect (Baldwin, 1896), Schmalhausen's ontogenetic stabilization (Schmalhausen, 1949), and Waddington's canalization (Waddington, 1953). Although a number of significant differences exist (Gilbert, 1994), all these proposals incorporate the common idea that phenotypic plasticity makes it possible the introduction of an adaptive trait that, after a few generations, would come to be assimilated by genes making thus possible its spread and maintenance over time. West-Eberhard's proposal on phenotypic accommodation is also meant to this idea: phenotypic plasticity provides adaptive variations that tend to increase fitness and phenotypic variations are then canalized to the genetic level due to preexistent *cryptic genetic variants* (Kirschner & Gerhart, 2008). The core idea is then that cells have enough unused genetic material available that can be put at

work to stabilize the developmental pathway which produces the phenotypic variant. This evolutionary process is quite different from the one envisaged by the MS, because variants originate during development. Thereafter, they are preserved by non-genetic sources of inheritance (cf. Section 4.2.1), and are eventually stabilized at the genetic level. In this view, no genetic changes take place: "plasticity is the basis of phenotypic change in the absence of genotypic change" (Bateson & Gluckman, 2011, 100). It is phenotypic evolution without genotypic evolution. As West-Eberhard (2003, 29) put it, "genes are usually followers, not leaders, in evolutionary change".

The recognition of phenotypic plasticity is not new, nor is the notion of NoR, which was introduced by Richard Woltereck in the early 20th century (Woltereck, 1909; Sarkar, 1999, 2006 for a historical overview). However, the role of phenotypic plasticity, NoR, and other related processes such as the Baldwin Effect, stabilizing selection, and canalization were disregarded by the MS (cf., for example, Dobzhansky, 1955; Simpson, 1953). The reason for dismissing these phenomena is pretty clear: if they were interpreted as their proposers suggested, they would fall outside the scope of the MS. So an alternative interpretation was needed. This alternative consisted in seeing phenotypic plasticity as arising from prior natural selection processes acting on gene pools. It was not about organisms adaptively responding to environmental conditions, but about previous evolutionary processes selecting genes capable of producing different traits in different environmental contexts. Thus, the Baldwin Effect, stabilizing selection, and canalization were subsumed by the classical framework of populational biology. Organismal causation was transferred to populational causes. Interestingly, those contemporary defenders of the MS that do not see phenotypic plasticity as a problem for the MS also adopt the same explanatory strategy: to understand phenotypic plasticity as the result of natural selection —i.e. to explain phenotypic plasticity in terms of populational biology.

Having said so, however, we are already in possession of some clues suggesting why the genetic interpretation of phenotypic plasticity does not fit with the one provided by the DT (see, for example, Sultan, 2019). Even if it were really the case that genes encrypt possible developmental outcomes, this would only be possible if genes' actions took place within the cell. But, as soon as the logic of the process is inverted —i.e. it is cells that control genes and not the other way around, the expression of the coded protein products (not traits) from cryptic genetic variability is only possible in the context of organismal development and it is related to the needs and possibilities of the organism. Transferring organismal causes to the populational level is a consequence of misunderstanding development and overestimating the role of genes in development.

Niche construction

The idea of niche construction was originally suggested by Richard Lewontin (Lewontin, 1983a), although it should also be recognized the contribution by West and King (1987). Niche Construction Theory, however, was only later developed under its current name by the likes of John Odling-Smee, Kevin Laland and Marcus Feldman (Laland, Matthews, & Feldman, 2016; Laland, Odling-Smee, & Feldman, 2019; Odling-Smee et al., 2003; Okasha, 2005; Scott-Phillips et al., 2014). Lewontin's idea of construction is indeed an antagonist to the idea of adaptation. Instead of organisms passively confronting their environmental conditions, organism are presented as actively constructing their niche. From this perspective, an avenue is opened to try to understand the different organismal activities that change the rules of the game. Instead of waiting for a random phenotypic variation that fits the environmental problem, the organism has the power to change the problem and make its trait fit with the new environment.

Since its original formulation at the beginning of the 21st century, niche construction was presented as "the neglected process in evolution", which was the subtitle Odling-Smee, Laland and Feldman chose for their 2003 book. No wonder, since it openly challenges Explanatory Externalism with the claim that organisms also introduce an adaptive bias in the processes of selection. Since then, a number of debates ensued, such as the one published in *Nature News* around the extent to which the MS is actually capable of incorporating niche construction theory within its framework or, instead, a radical modification of the standard evolutionary theory is needed (Laland et al., 2014). Indeed, the debate has been passionate, often acrimonious, and it is still ongoing —see Laland, Odling-Smee, and Endler (2017); Laland and Sterelny (2007) for further references and arguments from the reformist side and Futuyma (2017); Gupta et al. (2017) as representatives of the other contenders. So just as in the case of phenotypic plasticity, different positions are at play here.

It will be useful to distinguish between two different notions of niche construction. The first notion is the one originally sketched by Lewontin —retaken later by, for instance, Stotz (2017)— which referes to individual-level processes of construction: organisms altering their environmental conditions to increase their fitness. This notion of fitness is Darwinian/individual. In contradistinction, most proponents of niche construction tend to favor a populational view (Laland et al., 2016, 2019; Odling-Smee et al., 2003; Scott-Phillips et al., 2014). This fact is made explicit when they also claim that niche construction is an *evolutionary process*. According to this view, niche construction has the power of modifying the fitness of populations.

I believe that the main challenge to the MS arises by taking *individual niche construction* as the relevant phenomenon at play. This is not to say that the

populational notion should be abandoned, it just means that the individual-level outlook better captures the core of the DT. Indeed, when contemporary defenders of the MS evaluate the adequacy of the challenges presented by niche construction processes, they usually discuss the populational notion of niche constructions (e.g. Futuyma (2017); Gupta et al. (2017)). From this populational perspective, it may be expected that niche construction can be explained within the framework of the MS. The explanatory strategy is similar to the case of phenotypic plasticity: to argue that niche construction processes are just a slight modification of standard populational genetics; populational niche construction may be incorporated within the mathematical models of populational biology, as (Gupta et al., 2017) argued.

The principal reason for adopting an individual niche Constitution conception, therefore, is that the real challenge to Explanatory Externalism comes from taking organisms as explanatory necessary in evolution. The mechanisms of development are a crucial source of heritable phenotypic variations. The individual interpretation of niche construction clearly captures this role of organisms: individual mechanisms during ontogeny contribute to changing fitness values. However, the notion of populational niche construction can be defined without taking into account such individual-level mechanisms. It is just an additional input to the populational dynamics of selection. Indeed, those who do not see a serious challenge in the idea of niche construction believe that its accommodation within the MS requires just a minimal modification to the standard framework of population genetics (Gupta et al., 2017). I am not going to take sides in this since it is not relevant for my project to assess the actual revolutionary spirit of the populational interpretation of niche construction. I will just limit myself to suggest that such revolutionary vibes are definitively present once we advocate for an individual-level, Lewontin-style interpretation of niche construction. Having said so, by the end of this chapter I will make explicit how the tension between the DT and the MS arises when these two interpretations —the individual and the populational— are mixed in connection with niche construction.

During ontogenesis, organisms can do many different things to alter their external conditions. One obvious way to do this is to move and go elsewhere: migrate to another ecosystem (which also includes cases of species invasion). Another way is to actively construct the niche the organism will live in adjusted to the organism's life conditions. Finally, organisms can also modify their external conditions by changing the environmental circumstances they live in. These are different sorts of *material niche construction*: different ways in which the environmental scenario is modified by the active participation of organisms during ontogenesis. In Section 4.1.3 I will introduce the notion of *experiential niche construction*, which also refers to an individual-level phenomenon which does not involve external, material processes of niche construction but internal processes of construction.

Self-organization

The role of internal self-organization in development as a source of adaptive evolution has been one of the classical bones of contention within developmental biology at least since Aristotle's original observations and, for example, the source of the historical debate between epigeneticists and preformationists (Gould, 1977; Pinto-Correia, 1997). It was well recognized by Kant and, since, self-organization became one of the theoretical cornerstones of the organicist framework pursued by pre-Darwinian embryologists (Gilbert & Sarkar, 2000; Lenoir, 1989). In the 20th century, the importance of an organismic level of analysis figured prominently in the thought of the members of the Theoretical Biology Club (cf. Section 6.2 and Nicholson & Gawne, 2015; Peterson, 2017 for historical overviews); it was also at the core of the first wave of the cybernetics movement (Rosenblueth, Wiener, & Bigelow, 1943; Wiener, 1948). Even though the role of self-organization and its implications for the emergence of novel phenotypic traits was well recognized and studied, it was, as expected, put aside from the evolutionary theory of the MS. The inner and global dynamics involved in self-organization as sources of adaptive variation stand in clear opposition to the Explanatory Externalism of the MS. As Edelmann and Denton (2006, 579) explain, "[b]iological self-organization [...] is a fundamentally different means of generating complexity [...] self-organization may be therefore considered a complementary mechanism to natural selection as a causal agency in the evolution of life". In the last decades, the role of self-organization as a distinctive evolutionary force has acquired new momentum through the work of different scholars in a variety of fields (Camazine et al., 2003; Goodwin, 1994; Kauffman, 1993, 1995, 2000, 2019; Maturana & Varela, 1980; Müller & Newman, 2003; Reid, 2007; Salthe, 1993, for some illustrative examples).

The significance of an organismic level of analysis involving the dynamics of the whole system becomes apparent once we attempt to tackle the distinctive properties of living beings. There are two interconnected peculiarities concerning the dynamical organization of living beings which will be central in Chapter 6. One is *thermodynamical openness*. Thermodynamical open systems —also known as dissipative systems— live in a far-from-equilibrium thermodynamical state, or, to use Stuart Kauffman's expression, "at the edge of chaos". This means that they are constantly exchanging matter and energy with the environment to preserve their inner structure and functionality. The clearer case of this is cellular metabolism, where the cell must be all the time interacting with the environment through the membrane to obtain the work needed to preserve its functionality, organization, reproduction, and repair the system. The other feature is *operational closure* (Maturana & Varela, 1980; Mossio, Montévil, & Longo, 2016). It concerns the inner organization of the system that separates it from the environment. Organisms form coherent wholes by the dynamical interconnection among their parts that ensure

the preservation of their autonomy (Moreno & Mossio, 2015).

These distinctive features of organisms are central in evolutionary theory for two reasons: for their role in a theory of variation, and for their role in a theory of emergent properties. Concerning the first point, the renewed interest in self-organization lies in its capacity to account for the construction of new phenotypic traits as the effect of the interaction of the systems' components (cf. Section 5.1.1 for Stuart Newman's work on inherent emerging patterns in living systems). As Goodwin (1994) argued, many properties of living beings arise as a consequence of the physical interaction of the organism. Moreover, self-organization is crucial for organisms to regulate their life conditions. Many changes or malfunctions in one part of the system can be repaired by changing other parts of the system. The rise of new traits must be analyzed in the context of the whole system and its environmental coupling. Self-organization is thus presented as an alternative view of variation beyond the gene-based account defended by the MS.

The idea of emergence is also connected with the organicist foundations. Accordingly, emergent properties are those that are not present in any of the parts of the system, but once such parts interact in a particular way, emergent properties arise at the level of the system. Although they are a source of intense theoretical debate, emergent properties in life science could correspond to the notion of life itself, the idea of organismal agency, or the emergence of psychological categories. In these cases, emergence exists because the whole system has a property that is not present in any of its parts: no molecule is alive but certain organizations of molecules are; no neuron has beliefs, but some organizations of neurons do. Here the slogan is *more is different* (Anderson, 1972): the interaction of parts generates a difference between the intrinsic properties of the parts and the properties of the whole that are observed in the interactions. In this sense, self-organization is central to understanding the emergent properties of life, properties that are not present in non-living systems, such as the physicochemical entities of which living beings are made of. As expected, such emergent properties cannot be given a bottom-up explanation: we cannot posit a sub-organismal entity —genes, for example— appropriate for accounting for the emergent properties. In this context, emergent properties and self-organization as a source of adaptive complexity suggest a different answer to Schrödinger's question than the one offered by the gene-centrism of the MS. Instead of assuming that the order of living beings comes from prior ordered structures, as Aristotle had already recognized self-organization suggests an order-from-disorder strategy (Saetzler, Sonnenschein, & Soto, 2011). Order arises as soon as disordered entities interact.

Self-organization and the order-from-disorder strategy will be the central topic of Chapter 6. By now, it is enough to highlight that the self-organizing and emergent properties of living beings defy the atomistic view of complexity fostered

by the MS. This view came from Mendel's experiments and Weismann's views on inheritance. As noted, in Mendel's experiments, discrete and distinctive phenotypic outcomes are connected to discrete and distinctive —atomic— inherited units. Weismann's barrier also meant that the locus of complexity must precede development and be already present in the inherited units of the germ cells. This view then pictures adaptations in a LEGO-like format: evolutionary histories acting on discrete and distinct traits —atoms— produce adaptive complexity that is maintained and spread through atomic inherited units. As Edelmann and Denton (2006, 587) remarked, "Neo-Darwinian complexity necessitates self-specification by the genes, not self-organization by the products of the genes!". In contrast, self-organization stresses the epigenetic (non-preformed) and interactive nature of adaptive complexity. Self-organization is not coded in any part of the organized system but emerges from wholeness. No adaptive complexity arising by self-organization could be explained in an atomistic, LEGO-like way because it hinges on the dynamical interaction of the system's parts.

Moreover, self-organization and emergence are distinctive marks of living beings. Certainly, other systems show self-organizing properties, such as Bénard Cells, tornadoes, or some artificial systems. Yet none of them exhibits the same qualities present in living beings regarding their thermodynamical openness and operational closures: "emergence resides in the primary qualities of life: simple persistence through reproduction, self-maintenance, and self-organization" (Reid, 2007, 394). This point could be taken as a clue when seeking an explanation of aptness. As I have already pointed out elsewhere in this thesis, the explanation of aptness is the core reason for the quest after legitimate teleological explanations in biology. Teleological explanations are a distinctive feature of the life sciences. Once this fact is recognized, as Kant did, the role of self-organization becomes central in the search for natural teleology. In Part III I will revisit with more detail this idea in order to pinpoint the connection between emergence, self-organization, and teleological explanations.

The evolutionary role of self-organization, as well as that of phenotypic plasticity and niche construction, is therefore connected with the opposing stances concerning the actual working of natural selection. At one extreme, as explained in Section 2.2.3, we have the positivist view, which sees natural selection as an active process for creating adaptive complexity. At the other extreme, the negative view understands that "Natural selection cannot explain the origin of new variants and adaptations, only their spread" (Endler, 2020). In this sense, natural selection operates as a filter and a conservative process. It ensures that those successful trials are repeated in future generations while eliminating the misguided ones. As Reid (2007, 27) wondered, "if natural selection is the filter, what's making the coffee", that is, if selection is the process of filtering, another mechanism

must be responsible for producing the materials natural selection needs to choose from. According to the negative view of natural selection, these mechanisms have something to do with individual development and self-organization; there lies the creative character of nature:"[s]elf-organizing material patterns may be selected by, but *not created by natural selection*" (Edelmann & Denton, 2006, 598; emphasis in the original).[2]

4.1.3 The organism determines what is relevant

The previous two subsections were devoted to introduce a number of research areas that in the last decades have been yielding a considerable amount of output that makes it possible to articulate a challenge of the MS on the basis solid experimental and theoretical evidence. This subsection is mostly conceptual, as it analyzes the relation between niches and organisms posited by Explanatory Externalism. The ideas expounded here are almost entirely based on Richard Lewontin's work (Levins & Lewontin, 1985; Lewontin, 1974a, 1974b, 1978, 1983a, 1983b, 2000; Lewontin & Levins, 2007).

In Section 2.2.1 I introduced Lewontin's lock-and-key metaphor on adaptation. The metaphor is a particularly graphic way to illustrate how the Explanatory Externalism of the MS thinks about niches. As noted, environmental problems are the only source of adaptive bias in evolution. Confronted with such problems, organisms cannot propose some adaptive solution. It is just a matter of waiting for a random variation to occur that is capable of bringing about the solution that fits the organism. Niches, in this picture, pre-exist and are separated from the organism. Organisms are passive objects in the course of adaptive evolution. The following quotation offers a perfect summary of these ideas:

> According to Darwinism, there are mechanisms entirely internal to organisms that cause them to vary one from another in their heritable characteristics. In modern terms, these are mutations of the genes that control development. These variations are not induced by the environment but are produced at random with respect to the exigencies of the outside world. Quite independently, there is an outside world constructed by autonomous forces outside the influence of the organism itself that set the conditions for the species' survival and reproduction. The inside and outside confront each other only through the selective process of differential survival and reproduction of those organic forms

[2]The negative-positive debate and its connection with phenotypic variation is an old one. In the 19th Mivart (1871) defended the negative view which motivated Darwin to recognize the issue of variation as an unresolved element in his theory; cf. Moczek (2008) for a contemporary exposition of the issues.

> that best match *by chance* the autonomous external world. Those that match survive and reproduce, the rest are cast off. Many are called but few are chosen. (Lewontin & Levins, 2007, 230; emphasis in the original)

The previous two subsections helped us to see that the organism is far from being a passive entity. The ontogenetic trajectory and organizational dynamics of each system are coupled with its living conditions in a way that the organism' activity is adaptively directed to them. Once the complexity of living systems is highlighted as a crucial source of adaptive evolution, and particularly, of phenotypic variation, "the organism cannot be regarded as simply the passive object of autonomous internal and external forces; it is also the subject of its own evolution" (Levins & Lewontin, 1985, 89). Lewontin pictures organisms as subjects rather than as objects (Godfrey-Smith, 2017; Walsh, 2018). His work hosts most of the original theoretical insights that feed current work in the area of organismal agency, a topic that will come to the center stage in Part III. As we will see presently, it is also a challenge to replicator agency defended by Dawkins and other advocates of gene-centrism. Lewontin's ideas must be kept in mind when discussing the connection between organismal agency and evolutionary theory.

Moreover, the non-passivity of organisms requires niches to be defined in relation to the organism. It is not anymore a question of the environment proposing riddles. Life conditions are produced by the organism itself through different processes of niche construction. Organisms and niches are (dialectically) co-constructed: the organism' properties determine which environmental properties are *relevant* (Lewontin, 2000) for it, while the environment is crucial to determine the boundaries of living beings, in a way that "just as there is no organism without an environment, so there is no environment without an organism" (Levins & Lewontin, 1985, 99). There is no point in asking for the environmental context of an organism without knowing the organism we are talking about: "[i]f one wants to know what the environment of an organism is, one must ask the organism" (Lewontin, 2000, 54).

In the previous subsection, I introduced the notion of niche construction. I also introduced the distinction between an individual interpretation from a populational one. The former refers to ontogenetic processes, and it is the one defended by Lewontin. However, as, for instance, Chiu (2019) suggests, we can distinguish different niche construction processes at the individual level. Lewontin himself acknowledges the two kinds of niche construction processes that take place during ontogenesis (Godfrey-Smith, 2017): one process refers to material changes in the environment, and the other concerns the internal construction made by the organism about those features of the word that constitute its niches. So one process refers to extrinsic material changes, while the other concerns intrinsic construc-

tion of what is the environment of an organism. The latter idea is connected to Lewontin's dictum that "the organism determines what is relevant" (Lewontin, 2000). This determination brought about by the organism is not only about those material things of the environment that the organism needs. It is also about the boundaries of its own reality. The organism is connected only to those things that somehow —perceptually, physically, or otherwise— affect its system. The notion of *experiential niche construction* refers to the second notion of niche construction (Aaby & Desmond, 2021; Heras-Escribano & De Jesus, 2018; Sultan, 2015). "The concept of niche construction", as Sultan (2015, 37; emphasis added) claims, "can be further extended to include phenotypic adjustments that permit the organism to *experience* a given set of conditions as more favorable, without either changing those conditions or moving to different ones". There, construction is also not conceived of as an external process of changing the niche. It is presented as the process through which the organism constructs what is real for it on the basis of its experience and that is enabled by signaling systems[3]

4.2 Beyond Replication

While the previous section challenged Explanatory Externalism, this one focuses on Replicator Biology. I will start by distinguishing different kinds of questions concerning inheritance: a material question and a conceptual one. This taxonomy of questions at play will help to better identify which theses of the MS are challenged. I will continue in Section 4.2.1 by introducing extended inheritance systems as a response to the material question: the reduction of inheritance to the genetic level advanced by the MS is empirically flawed. In Section 4.2.2 I will present a developmental conception of inheritance as a response to the conceptual question: the view of inheritance as a process of replication ignores the causes that produce cross-generational resemblance.

Recall that, previously, in Section 4.1.1, while I was presenting the post-genomic era, I posed two questions: a *material* question and a *conceptual* one. The material question concerned to what an extent does the molecular gene capture what genes really are and how they function. I argued that the reactive genome has come to replace molecular genes. The conceptual questions was different and to a certain extent independent of the material question. It had to do with whether the explanatory status of genes deserves to be changed or whether the post-genomic gene still plays the explanatory role that the MS intends. If the latter is the case, then the post-genomic world would still be an MS world. It would just be a step forward towards a better understanding of the nature of genes, as

[3]The role of singling systems in developmental regulation and the construction of "a point of view on the world" (Godfrey-Smith, 2017, 4) of each organism will be explored in Chapter 8.

Watson and Crick's discovery was. In Section 4.1.1 I argued that post-genomics also calls for a re-conceptualization of the explanatory role of genes. Particularly, the misconception turns around treating the gene and not the organism as the proper unit of developmental analysis.

This section is about different challenges to Replicator Biology as the view of inheritance forged by the MS. As noted above, here it is also relevant to distinguish between two questions:

1. **Material question:** Is inheritance exclusively genetic?
    - From genes as the only source of inheritance to multiple, extended systems of inheritance.
2. **Conceptual question:** Does Replicator Biology preserve its explanatory status?
    - From a Replicator Theory of Inheritance to a Developmental conception of Inheritance.

In this case too, the question about the material bases of inheritance is different from the question how inheritance should be conceptualized. Keeping these two questions apart is relevant because many scholars accept extended inheritance systems but still promote a replicator view of inheritance. Therefore, the fact that one accepts the existence of extended inheritance systems does not necessarily mean that one is also interested in explaining the causal role of such systems in development. The challenge to the MS does not reduce to accepting non-genetic inheritance systems but it also entails challenging the replicator view initially forged by Weismann.

I will now move to introduce extended inheritance systems and their importance in evolutionary theory. Extended Inheritance systems represent an answer to the material question about inheritance (Item 1). In Section 4.2.2 I will deal with the conceptual question of inheritance (Item 2). There, I will argue that the shift from genes to organisms as the units of developmental control, as explained in Section 4.1.1, also demands a developmental theory of inheritance.

4.2.1 Extended inheritance systems

The importance and consequences of extended inheritance systems were largely enunciated by geneticist Eva Jablonka and colleagues (Avital & Jablonka, 2000; Gissis & Jablonka, 2011; Jablonka, 2007; Jablonka & Lamb, 1995, 2014, 2020; Jablonka & Raz, 2009). Jablonka strongly defends the existence and importance of extended inheritance systems as well as the need to rethink the very conception of

inheritance bequeathed from 20th century biology. In this subsection, I will present the four kinds of inheritance systems proposed by Jablonka and her collaborators.

The first one is the genetic system of inheritance. It possesses two distinct qualities. First, it is reliable: genetic transmission among generations is a stable process; the genetic properties of parents are transferred with high fidelity. The reliability of inheritance is central to the stability of evolutionary processes and the possibility of cumulative selection. As already noted, if new generations would not resemble their ancestors, it would be difficult to understand how populations, species, and other taxa could emerge. Moreover, the reliability of genetic transmission allows selection to operate across many generations to bias populations by accumulating selected heritable traits over time. Certainly, the genetic system of inheritance has unique and important properties for evolutionary processes. However, it is not the only system of inheritance present in nature. This remark is the clue to understand that the role of inheritance in evolution is not just about stability and preservation across large time scales.

The second system of inheritance is the epigenetic system. In this context, 'epigenetic' is used in a narrow sense and referring the processes of protein expression in cell development, as it was introduce by Nanney (1958), and not in the broader sense concerning any process involved in the mapping of genes to phenotypes, as it was used, for example, by Waddington (1941) (cf. Stotz & Griffiths, 2016, for different conceptualizations of epigenetics). In the broad sense, the one employed by Waddington, epigenesis refers to the process of trait construction —the map from genotypes to phenotypes. This sense includes all non-genetic modes of inheritance. In a narrow sense, epigenetics refers to those processes concerning cell formation, division, reproduction, and death. This sense of epigenetics captures Jablonka and Lamb's notion of cellular modes of inheritance. These cases involve many relevant and interesting phenomena, such as histone modification, chromatin marking, self-sustaining metabolic loops, structural forms of cellular inheritance, heritable RNA variations, among other things (cf., Jablonka & Lamb, 2014, for a detailed account of cellular/epigenetic systems of inheritance).

This system, as the previous one, is ubiquitous in nature. The particularities of these modes of inheritance are twofold. Firstly, they are central in developmental processes. They are not merely a material basis used to express genetic information. Rather, they constitute an important source of developmental specificity. In other words, changes in these resources entail modifications both in developmental outcomes and in inherited material. Secondly, they interact with genetic material in a bi-directional way. This, as previously noted (Section 4.1.1), means that genes are also epigenetically controlled, and, therefore, that their functions cannot be assessed outside of the context of the cell. In this sense, genetic and cellular modes of inheritance must work together in development and evolution in order to both

produce new variants and keep them stable over generations.

The other two inheritance systems are the behavioral system of inheritance and the symbolic system of inheritance. The former is present in animals, while, according to Jablonka and colleagues, the latter is exclusive of humans. Behavioral and symbolic systems of inheritance are mediated by different sorts of learning processes, such as imitation, operational conditioning, or even by teaching. A particularity of these inheritance systems is that they are not exclusively present in parent-offspring relationships. They both allow for different sources of inheritance within a population. Learning processes could be mediated or scaffolded by groups, such as families or communities. Moreover, they could take place at different scales of generations. There are parent-offspring, grandparent-offspring, or even horizontal transmissions of behavioral and symbolic characters involving members of the same generation.

Extended inheritance systems are part of nature. Why should we neglect their evolutionary relevance? One argument is that they are not as spread across taxa as the genetic system of inheritance. While genetic systems of inheritance are present in any species and in any sort of inter-generational reproduction, epigenetic systems are just present in some species, thus they are a coda of evolution rather than a central element. This criticism could be answered, principally, on empirical grounds: epigenetic inheritance "has been found in all organisms in which it has been sought" (Jablonka & Lamb, 2020, 23). So while maybe the behavioral and symbolical system might be exclusive of some species, both the genetic and the epigenetic systems are ubiquitous in nature; it is an empirical fact that in all species the construction of genes is mediated by cellular processes that are epigenetically inherited (cf. Jablonka & Lamb, 2020, for further arguments and many examples).

Another possible reason for skepticism towards extended inheritance systems is that they are not reliable enough to secure evolutionary processes. Unlike genetic systems, epigenetic variations need not be maintained through many generations or passed on with the reliability that genes are. However, as Jablonka and Lamb (2020, Ch. 4) argue, the importance of an inheritance system is relative to a time scale. If one is interested in how variations could be safely maintained across millennia, then genes are the place to look at. Nonetheless, extended inheritance systems come to the fore at shorter time scales. *Inheritance needs not to be a long-term phenomenon to change the evolutionary scenario.* Extended inheritance systems become central in the epigenetic origin of phenotypic variation. If adaptive phenotypic variations arising during epigenesis —that is, as changes in developmental systems— do not have the chance to be inherited, the role of organisms in producing new variants (cf. Section 4.1.2) cannot be connected to the evolutionary arena. The importance of extended inheritance systems is tied to the role of organisms in evolution. This idea was already present in the definition of such

processes as canalization, stabilizing selection, and phenotypic accommodation. Once variations arise, time is needed for them to become genetically fixed; this necessary amount of time is supplied by non-standard modes of inheritance. So extended inheritance is a fundamental ingredient in the DT, and it is best appreciated in the origins and stabilization of variations, rather than in its maintenance and distribution.

Once the role of extended inheritance systems in evolutionary processes is recognized, what are the consequences of this for the MS? The first one is that development and inheritance cannot be kept separated in time: not everything is about donation at conception (Mameli, 2005). Inheritance also takes place during development. In this sense, epigenetic changes during development can give rise to new heritable variation. Moreover, extended systems of inheritance differ from the genetic system insofar as variation in the former is usually adaptively directed, while variation in the latter usually is not. That is, the sources of most genetic variation is a non-adaptively directed, random processes, such as drift or mutation. In contradistinction, epigenetic modifications take place within a developmental context. Such variation occurs as an organismal answer to the conditions of life. In the previous section, I already identified a number of processes that give rise to adaptive phenotypic variation, such as plasticity, niche construction, and self-organization. Now we can add that such variation may be transferred down across generations. So extended inheritance is crucial for understanding the origin and maintenance of adaptive phenotypic variation.

This view incorporates several ingredients that make it sound like a neo-Lamarckism of sorts, and it is certainly the case that sometimes it is so presented (Jablonka & Lamb, 1995). Neo-Lamarckism, or this particular version of neo-Lamarckism at least, is nonetheless still Darwinian: " 'Lamarckian' inheritance would not exclude Darwinian selection. It would complement it, providing yet another source of diversity" (Noble, 2006, 95). As it is well known, Darwin himself accepted Lamarck's theory of inheritance, so, in principle, there is nothing anti-Darwinian in Lamarckian inheritance. To be sure, above I sketched an evolutionary process taking place at the epigenetic level: phenotypic *variation* mediated by organismal responses during development, directed towards an adaptive state that secures its increment on *fitness* and which are passed on through extended *inheritance* systems. However, while this Lamarckian dimension is Darwinian, it is definitively not neo-Darwinian. Mostly because development and inheritance are not fractionated. Or, more specifically, first because inheritance is not depleted at conception but is present throughout ontogeny, and second because somatic changes could be transferred down in generations. In other words, epigenetic processes are also central for explaining the resemblance within a lineage and variation on such epigenetic processes may entail variation in further generations. So it is

relevant to note that extended inheritance systems are entirely compatible with Darwinian evolution, just as Lamarck's theory is not necessarily in opposition with but complementary to Darwin's theory. The problematic step was taken by the MS and its commitment with genetic systems of inheritance. As Eva Jablonka and Marion Lamb put it:

> Although the current gene-centered version of Darwinism —*neo*-Darwinism— is incompatible with Lamarckism, Darwinism is not. In the past, Lamarckism and Darwinism were not always seen as alternatives: they were recognized as being perfectly compatible and complementary. In the light of epigenetics, they still are. Recognizing the role of epigenetic systems in evolution will allow a more comprehensive and powerful Darwinian theory to be constructed, one that integrates development and evolution more closely. (Jablonka & Lamb, 2002, 95)

4.2.2 Extending inheritance: beyond replication

Extending inheritance within the replicator framework

As it was explained at the beginning of this section, it is essential to appreciate that challenging Replicator Biology is not about just extending inheritance systems. Extension certainly has an important consequence, already pointed out above (Section 4.2.1). But the extension of inheritance systems could still be integrated within a replicator framework. We can appreciate the consistency between extended inheritance and Replicator Biology just by looking into different replicator proposals that integrate non-genetic modes of inheritance. This fact was clear when considering Richard Dawkins's work, who sees the concept of replication as an essential ingredient of *all* evolving systems. In the living world, according to Dawkins, the role of replicator is played by the gene, but his definition of (genetic) replicator does not necessarily match that of the molecular gene (Dawkins, 1982, 81–85). Indeed, Dawkins has proposed that there are non-genetic units of replication, namely memes, and memes, qua replicators, are the units of cultural inheritance (Dawkins, 1976). So, it is clearly the case that certain forms of extended inheritance do not necessarily stand in opposition to replicator ideas, as witnessed, for example, by the different models of biological and cultural coevolution developed since the late 1970s and early 1980s (Boyd & Richerson, 1985; Durham, 1991; Lumsden & Wilson, 1981; Richerson & Boyd, 1978, 2005).[4] Other

[4]These models are also a good example to show that adherence to the idea of (faithful) replication is also not dependent on the idea of particulate inheritance. Thus, for example, the dual-inheritance theory developed by Robert Boyd and Peter Richerson rejects the idea that Dawkins's and Durham's memes and Lumsden and Wilson's culturgens are adequate as units

scholars have also held ideas along these lines. For example, Sterelny, Smith, and Dickison (1996)'s Extended Replicator is just an extension of the gene-version of the replicator theory to include some epigenetic effects. Similarly, as it will be discussed later on, Shea's theory of inherited representation is also a replicator stance that accepts extended inheritance systems. To be sure, as pointed out by Griesemer (2000a, 348), even some of the original presentations of Developmental Systems Theory —e.g. Griffiths and Gray (1994)— appealed to the notion of replication at the level of life-cycles, which did not mean a breach of the replicator scheme but "only a nominal reunification of heredity and development".

As shown by Ron Amundson in his wonderful *The Changing Role of Embryo in Evolutionary Thought* (Amundson, 2005) the concept of inheritance was borrowed by biology from social science to describe parent-offspring interactions (see also Keller, 2010, 21). The notion was fully integrated in biological thought in the 19$^{\text{th}}$ century, but not necessarily as a process independent from development as many see it today. I already explained how the genetic theory of inheritance was constructed by the likes of Mendel, Weismann, and later on Morgan, and how Replicator Biology grew up in this scientific niche. But inheritance was not always separated from development. This divorce was brought about by the MS. Before Weismann, in Darwin's or Lamarck's times, for instance, inheritance was seen as part of development. As Amundson notes, in pre-Weismannian views, inheritance was seen as "the production of parent-offspring similarities, and this production [took] place throughout epigenesis. Heredity [was] an epigenetic process [...] the causes of *heredity* [were] exactly the same as those of *development*" (Amundson, 2005, 142-143). In this sense, to argue for a developmental conception of inheritance against a replicator view is tantamount to rescuing old views about inheritance. This idea has also been defended by Jablonka and Lamb, who contend that the experimental advances in Extended Inheritance Systems need to be accompanied by a reconceptualization of the very idea of inheritance: "[w]e need to return to an earlier, development- and organism-oriented view" (Jablonka & Lamb, 2020, 1), a view where "[h]eredity is seen as an aspect of development, and the origin of heritable variations and their transfer are therefore analysed as developmental processes" (Jablonka & Lamb, 2020, 55).

So far, I have argued that the answer of the MS to the material question must be revised: inheritance is not exclusively genetic. Recall, however, that the material questions are different from the conceptual ones. My aim now is to explain why the DT also calls for a new answer to the conceptual question about inheritance; i.e. why we should take a step back and return to pre-Darwinian conceptions of inheritance.

---

of cultural transmission, while sticking to the idea of replication as the basic process for the reproduction of such units; cf. Sperber (1996) and Richerson and Boyd (2005) for discussion.

From replication to construction

In Chapter 2, I observed that the theory of inheritance set up by the MS must be suitable for populational explanations. If evolution is possible, traits must be preserved among generations for selection to act on them. What is needed therefore is that selected traits be preserved across generations for cumulative selection to be possible. Replicator Biology conforms to such requirements: if inheritance is about the transmission of traits by genetic (or non-genetic) means, then the presence of traits across generations is a consequence of the stability of the transmission of traits. However, in doing so, Replicator Biology posits that the reappearance of traits in each generation is possible insofar as the information needed for constructing the trait is already present before development takes place. In other words, the resemblance of traits across generations is not a consequence of developmental processes but a matter of inherited information. This idea, even though seldom presented in informational terms,[5], was already present in Weismann. Wiesmann's experiments led him to conclude that epigenetic processes cannot alter inherited materials, such that whatever explains the resemblance of traits must be indifferent to epigenetic processes. This preformationism of sorts is at the core of Replicator Biology; i.e. the idea that developmental processes may be ignored because traits are replicated one generation after the other on the basis of inherited materials.

Previously, I dissociated the material and conceptual questions. However, now it is also relevant to point out the connection that exists between the *conceptual question about genes and the conceptual question about inheritance*. This connection is easy to appreciate if we look at the explanatory roles of genes and of replicator accounts in evolutionary theory. In both cases, we find the idea that the information needed for producing traits predates and is irrelevant to developmental processes. If this idea is well-supported empirically, then there is still room for Replicator Biology and for a treatment of genes as units of developmental control; i.e. the standard answers to the conceptual questions remain intact. But, if it can be shown that the information needed for producing traits does not predate development and developmental processes are not irrelevant for the explanation of traits construction, then we need new a answer to the conceptual question. So the common land is that the information for constructing traits predates and ignores developmental processes. Criticizing this idea lies at the core of an answer to the conceptual question.

The quest for a developmental theory of inheritance is motivated by the discovery of extended systems of inheritance, but above all by the renewed emphasis on the organism as the basic unit of development. Indeed, tacking the organism as the unit of development was also the key to answer the conceptual question

[5]Probably the first formulation of inheritance in informational terms is due to G. C. Williams (1966).

about genes. The organism itself is responsible for producing its own traits during development, and such process of construction is not specified anywhere before development takes place. Order does not preexist genes but comes into being during development. If the conceptual questions are connected, the answer to the conceptual question about genes (Section 4.1.1) is also connected with the conceptual question about inheritance.

At the core of any replicator view is that order preexists development: order comes from previously ordered matter (an order-from-order strategy). Conversely, a developmental conception of inheritance is aligned with the idea of the ontogeny of information: the construction of traits is not pre-specified anywhere. Order emerges from organismal regulation throughout development (an order-from-disorder strategy). So the reason why development cannot be specified by genes (whatever their material basis) is the same reason why development cannot be specified by replicator units (whatever their material bases): development is not a process of unfolding order, it is a process of creating order. In other words, the adaptive complexity of organisms does not precede the developmental processes that produce adaptive and complex organisms.

To systematize the ideas sketched so far, Table 4.3 on the next page complements Table 4.2 on page 89 to show two alternative ways of understanding the relationship between development and inheritance. The connection between the conceptual question about genes and the conceptual question about inheritance concerns, after all, the relation between inheritance and development. Understanding such a relationship is crucial for evolutionary theory: the rise of new species and taxa requires uniformity of traits in the individuals of a population. The MS explains this uniformity in a population by separating development from inheritance and positing pre-formed units of inheritance responsible for producing uniform traits across generations. In doing so, development becomes irrelevant insofar as development is not the source of order. The order of adaptive organisms preexists development. Alternatively, the DT has been pursuing a view where the uniformity of traits in a population is the consequence of developmental processes. In this sense, when can see both alternatives shown in Table 4.3 as two different ways of accounting for Darwin's missing element: the reappearance of traits across generations. Consequently, both alternatives fall under the umbrella of natural selection. In Chapter 2 I asserted that the MS represents a specific view about natural selection, while, as it should be clear by now, the DT endorses another conception of natural selection.

Modern Synthesis	Developmental Turn
Order-from-order	Order-from-disorder
Genes as the unit of development	The organism as the unit of development
Replicator conception of inheritance	Developmental conception of inheritance

Table 4.3: Alternative ways of explaining order: part II.

The abandonment of Replicator Biology also implies a backtrack of the path originally taken by the MS. One of the main consequences of replicator ideas was the displacement of organisms as second-class citizens in biology. Genes, not organisms, became the protagonists of life. Now, we need to take a step back and put the organism again at the center of biology. Paraphrasing Nicholson (2014), we can say that biology has traced a circle throughout its history: first it started as an organism-centered view, which was then replaced by a gene-centered stance, and now it is coming back to its original foundations. Once this movement will be complete, the distinction between replicators and vehicles will make no sense. Developing organisms do not serve as replicators.

Heritability in a developmental conception of inheritance

Let's conclude this section by noting an important point. I explained how, during the gestation of the MS, the concept of inheritance suffered two modifications. One was its reduction to the genetic level and its disconnection from developmental processes. This gave rise to Replicator Biology. The second one was the construction of a populational concept of inheritance incarnated in the technical concept of *heritability*. My goal now is to investigate whether challenging Replicator Biology entails a rejection of heritability. In other words, to what an extent Replicator Biology is necessary for populational explanations of evolutionary processes by natural selection?

It is useful to distinguish between two possible interpretations of the previous question. One interpretation concerns the adequacy of a non-replicator view of inheritance for the notion of heritability. A different interpretation is about the adequacy of a non-replicator view of inheritance for the view of heritability proposed by the biometric school at the beginning of the 20$^{\text{th}}$ century.

Regarding the first interpretation, I see no tensions between a developmental theory of inheritance and the notion of heritability used in populational explanations. As noted, heritability is a statistical measure concerning the degree of persistence of traits among generations. It says nothing about how such traits are

produced, but just averages those traits that reliably appear generation after generation. Developmental theories of inheritance certainly are interested about the mechanisms of trait construction. However, this is does not prevent them from taking statistical averages over the outcomes of developmental processes. This idea will be clearer once I introduce the statistical view of natural selection in the following section. Developmental theories focus on the reconstruction of traits in each generation, which is an individual-level phenomenon. Heritability, on the other hand, is the notion needed to deal with heritable traits at the level of populations. The absence of tensions lies precisely in the fact that these concepts refer to different sorts of processes —both in time and scale— involved in different kinds of explanations.

The second issue does deserve a critical analysis, however. The proposals of the biometric school seem not to fit with a developmental view of inheritance. That is clear, for instance, in Dawkins' recognition of the important role of Fisher's work in the gestation of the gene-eye view of inheritance (Ågren, 2021). Fisher's notion of heritability is based on two important ideas: (i) that the environment remains fixed across generations; and (ii) that heritable variations must be linked to genetic variations. None of these ideas are true in a developmental view of inheritance. Firstly, because the environment is a source of inheritance and is constantly changing during evolutionary processes, as the cases of niche construction illustrate, and secondly, because heritable variation cannot be linked exclusively to genetic variation, and we already know why.

Developmental theories of inheritance are therefore not necessarily incompatible with populational accounts based on heritability. Replicator Biology is not indispensable to understanding populational changes. However, the specific interpretation promoted by the founders of population genetics does require some relevant modifications.

4.3 The Statisticalist School

In the last decades, there has been an ongoing debate concerning the causal structure of natural selection, recently summarized in a critical work by Charles Pence (Pence, 2021). The main points of this dispute turn around the views of the so-called Causalist School against those of the so-called Statisticalist School. Even though there are many nuances within causalism, the core view was already presented in Section 2.2.3 under the idea of populational forces. According to the Causalist School, natural selection is a causal phenomenon taking place at the populational level and acting during evolutionary history. A classical stance within the Causalist School is Elliott Sober's, although, in the last decades, others scholars, such as Abrams (2012); Millstein (2006); Pence and Ramsey (2013); Ramsey

(2016); Reisman and Forber (2005) and Stephens (2004), have also taken sides in favor of the causalist position.

The statisticalist view was articulated and vindicated in 2002 by Walsh, Lewens, and Ariew (2002) and Matthen and Ariew (2002), and, since, other works have further developed its main tenets (Ariew, 2003; Ariew & Lewontin, 2004; Ariew, Rice, & Rohwer, 2015; Walsh, 2003, 2007b, 2019; Walsh, Ariew, & Matthen, 2017). In this section, I will briefly present it and point out the reasons why in my opinion it deserves serious consideration. The core idea is that individual causes, not populational ones, are the causes of adaptive evolution. In this sense, populational explanations of natural selection are the statistical consequence of individual causes. As Walsh (2019) presented it, evolution is a (populational) higher-order (statistical) effect of (individual) lower-level causes. I will focus first (Section 4.3.1) on how the Statisticalist School understands populational explanations and on the connection between the DT and the Statisticalist School. Next (Section 4.3.2), I will move to argue why and how the Causalist School may be taken as a non-Darwinian element within neo-Darwinism and how this connects with the division of explanatory labor propounded by the Statisticalist School.

4.3.1 Lower-level causes — Higher-order effects

The notion of fitness is the central one in any explanation by natural selection. Populations change due to fitness differences. The starting point in the statisticalist reading of natural selection is the difference between two notions of fitness: trait fitness and individual fitness. Trait fitness properly pertains to population thinking as the MS formulated it; trait fitness is about the fitness values of a population, about their trait types, not their trait tokens. In contrast, individual fitness —also named Darwinian fitness— refers to the fitness value of each individual. It is the Darwinian notion of fitness, and it concerns trait tokens, not trait types.

This difference lead to a central epistemological distinction, which refers to the different kinds of explanations involved in each notion of fitness. "Trait fitness is the average survivability of a group of individuals possessing a type of trait" (Ariew, 2003, 562), while individual fitness refers to the survival and reproductive capacities of an individual organism. The crucial difference is, therefore, that trait fitness is a statistical measure —a populational average— while individual fitness is causally assessed —it concerns the individual causal processes that determine a particular fitness for each individual. Trait fitness concerns the fitness value of a type which is measured by averaging over the individual fitness of trait tokens.

As explained, these two notions of fitness correspond to two different levels of analysis: the individual one and the populational one. Once this epistemological distinction is at place, the central thesis of Statisticalism is that all causes of evolution lie at the individual level: "[t]here is one level of causation; *all the*

causes of evolution are the causes of arrival and departure (the 'struggle for life') [...] It is 'proximate' causes all the way down" (Walsh, 2019, 238, 242; emphasis in the original).

The connection between the two levels therefore hinges on the connection between individual fitness and trait fitness. As defenders of the Statisticalist School propose, trait fitness is a statistical notion. It is obtained by looking into individual trait fitness and averaging them. Changes in individual fitness will produce changes in trait fitness. This idea is presented under the label of *analytic (mathematical) consequence* (Walsh, 2015) or, also, *statistical effect* (Walsh, 2007b). Trait fitness is a consequence of individual fitness. Such consequences are mathematically analyzed, with the tools of population biology developed since the beginnings of the 20$^{\text{th}}$ century. So what happens at the evolutionary level is a consequence of individual phenomena. Yet evolutionary processes are described in abstract terms involving the fitness values of trait types in a population. As Walsh et al. (2002) put it:

> In short, natural selection occurs only when the relative frequency of trait types changes in a population as a consequence of differences in the *average* fitness of individuals in different trait-classes. This is what we call the statistical interpretation of natural selection. (Walsh et al., 2002, 464; emphasis in the original)

As noted, the statisticalist view emphasizes the abstract character of populations in populational explanations. This concerns the capacity of looking at evolution in abstract terms without taking into account specific details about the causal processes that produce heritable variations on fitness. The issue of abstraction is important. Certainly, populations are not abstract entities. They are constituted by concrete individual organisms, but even if we now refrain from getting into the problem of determining the boundaries of species and groups, it is clear that this should not be an obstacle for treating populations as abstract entities in populational explanations. The abstract character of the population is at the core of the notion of trait fitness, as a statistical measure, in populational explanations. As the representation of populations is present in abstract, mathematical, statistical terms, the changes in populational structure —i.e. higher-order effects— are also defined in abstract terms. The study of natural selection processes driving populations to adaptations is a statistical analysis of the changes in populational structure due to changing rates in trait fitness. All in all, populations need not be abstract to be represented in abstract terms.

The difference between the Statisticalist and Causalist Schools is epistemological and concerns the nature of populational explanations. What are the reasons to support each view? Crucially, the debate between statisticalism and causalism is

connected with the issues presented in previous sections. Specifically, the emphasis on developmental causes of evolution pushes us to a statsticalist reading, while a replicator and externalist account make the individual level causally irrelevant.

No wonder that the Causalist School was born in connection with Replicator Biology and Explanatory Externalism, according to which populational forces distribute replicator units within populations in an adaptive-directive way (i.e. producing adaptations). Here, individual causation has no explanatory role in evolution. Natural selection impinges an adaptive bias on heritable outcomes according to their fitness differences, in such a way that how such differences come into being is pretty much irrelevant. If individual causes are out of the screen, we end by seeing natural selection as the driving force towards adaptive complexity.

But following the challenges of the DT, an alternative view on inheritance, variation, and fitness should be endorsed. As such, this biological theory does not yet exist. At least not in the same, fully articulated form as the MS. Some of the ideas were already introduced in this chapter. Other insights will be developed in the third part of this thesis. However, what brings together many non-classical biologists and philosophers of biology is the role of developing organisms in evolution. So the reasons for rethinking an alternative biological theory beyond replicator and externalist ideas are connected with the proper causes of adaptive evolution and, consequently, with the causalists vs. statisticalists debate. Having said so, the reemergence of individual causation in contemporary biology operates as a reason to defend the statisticalist reading over the causalist one. The biological backup of the Statisticalst School is, therefore, the particular understanding of the role of developing organisms in evolution driven by different areas in contemporary biology.

4.3.2 The division of explanatory labor

The Causalist School, statisticalists contend, is not Darwinian (Godfrey-Smith, 2009; Walsh, 2000, 2010, 2015). The gestation of the MS exemplifies a particular position within the many possibilities that evolutionary theory can take under the label of natural selection. In this journey, from Darwin's seminal idea to the MS's mature theory, lies the construction of populational forces driving populations to adaptive peaks or deserted valleys. As explained, the MS supported different tenets that fall within natural selection theory, but the MS is one among many positions that could be part of natural selection theory. In this sense, Darwin's view on natural selection is different from the MS view of natural selection. What is the difference?

The non-Darwinian character of the neo-Darwinian MS rests precisely on the difference noted by statisticalists. Darwinian fitness is individual fitness. MS fitness is trait fitness. Darwin's proposals have to do with how individuals struggle

for living longer and leaving more descendants. As noted in Chapter 2, the struggle for life should not be understood exclusively as competition between organisms, but as a confrontation to life in order to increase individual fitness. In other words, struggling for life encompasses all those processes carried out by the organism in order to stay adaptive and reproduce. Darwin's struggling causes individual fineness; Darwin's struggle for life is the locus of the causes of adaptive evolution. As Walsh stated, "[e]volution is adaptive because ontogeny is adaptive" (Walsh, 2007a, 195).

I already observed, on the basis of Godfrey-Smith's analysis (Godfrey-Smith, 2009), that classical formulations of natural selection eliminate any references to the struggle for life. Instead, they describe it in abstract terms, using statistical language, and avoiding any reference to individual causes. This fact illustrates that the locus of populational forces is not Darwin but the MS. Therefore, as Walsh contends, "[t]he source of the error [in the Causalist School], I believe, lies not in the Origin itself but in an erroneous metaphysical picture drawn from the Modern Synthesis theory of evolution. That theory explicitly construes selection as a force acting over populations of genes" (Walsh, 2000, 137). The result is that while Darwin's position is that the causes of adaptive evolution lie at the individual level, neo-Darwinians posited populational forces.

At the beginning of Chapter 2, I introduced Darwin's theory by highlighting that one of his main contributions was the importance of populational explanations to understand adaptive evolution. This gave rise to the so-called population thinking. The Darwinian character of neo-Darwinism is that both assume that populational explanations are needed for explaining the evolution of species. However, I also stressed that Darwin's view on the causes of adaptive evolution rests on individual causes not populational ones. So evolution is caused by individuals but explained by populations. How is this possible?

Denis Walsh labeled this conundrum the Paradox of Population Thinking. Notice, first, that in the context of the MS such a paradox does not even arise. Both evolutionary causes and evolutionary explanations are populational. The paradox, however, becomes apparent in the context of the two-force model (Walsh, 2003, 2019) which in turn arises because of the promotion of organismal causation. Notwithstanding, the two-force model is less radical than the statisticalist view, because, according to the former, evolutionary causes run both at the individual and the populational level. Natural selection (a populational phenomenon) is as much a cause of evolution as it is, for example, niche construction (an individual phenomenon). To the extent that the two-force model is committed to the individual causes of evolution, it must deal with the paradox of populational thinking.

The statisticalist view has an answer for this paradox (Walsh, 2019). The para-

dox is indeed diluted once we realize the division of explanatory labor in biology between the individual and populational levels that statisticalism promotes. We already encountered a similar sort of division of labor earlier: Mayr's separation between proximate and individual causes associated with two different levels of analysis and different explanatory roles in biology. However, the proposal here is different. The division of labor does not rest on a division of kinds of causes but on a division of explanatory strategies: the individual level provides causal explanations, while the populational one provides statistical explanations. The division of labor is central to stressing the non-reducibility and indispensability of any of the explanatory strategies. The notion of individual fitness is not sufficient to explain evolutionary events insofar as evolution is about the history of populations and individual fitness concerns an individual lifespan. Here is where trait fitness enters the scene. Crucially, trait fitness does refer to populational properties; it is a suitable notion to describe change in populations through time, something individual fitness cannot do. So statisticalism does not endorse the *ontogenetic fallacy* (Hochman, 2012) of attributing individual-level causes an explanatory role beyond its capacities.

In other words, we cannot understand evolution without statistical explanations, but evolution would lack a causal foundation without individual-level analysis. Without a populational dimension, evolution becomes development. Without an individual dimension, evolution loses its causal roots. In André Ariew's words:

> On my view evolutionary explanations are *statistical explanations of population-level phenomena* to be distinguished from 'proximate' or individual level causal explanations. The result is that evolutionary explanations are indispensable even if one knows the complete causal story about how each individual in a population lived and died. In other words, evolutionary explanations are not reducible to individual-level causal explanations. (Ariew, 2003, 561)

Everything considered, under the statisticalist reading the paradox appears not to exist. The fact that evolution is caused by individuals but explained by populations is not problematic for biology once we recognize the plurality of explanatory strategies and explanatory aims. As Walsh (2019) proposes, rather than the two-force model, the statisticalist view defends an epistemological distinction between two levels; i.e. statisticalism supports the *two-level model*. Accordingly, there are no two levels of competing forces trying to catch the biologists' attention. There is only one: individuals. However, two different explanatory levels are cooperating to help understand the complexity of the living world at different levels and across time.

In sum, the Statisticalist School provides a different interpretation of natural selection. This interpretation, however, is not disconnected from the ingredients of

natural selection. So a defense of statisticalism is tied to a *particular* view on the nature of natural selection, but without actually abandoning its place under the umbrella of natural selection. In other words, both schools accept that natural selection follows from inherited fitness differences, and both support Darwin's insight as well —i.e. that populational explanations are indispensable to understanding adaptive evolution. Consequently, the differences are not about the need for natural selection for explaining evolution; rather, the difference lies in the nature of explanations by natural selection.

4.4 Summary

At the opening of this chapter I already advanced that my aims here were twofold: to challenge the MS and to introduce an alternative evolutionary theory. In pursuing the first aim I laid bare several challenges to the pillars underpinning the MS. Explanatory Externalism was confronted with the central explanatory role of organisms both in development and evolution. Replicator Biology was confronted with extended sources of inheritance in addition to a number of theoretical challenges concerning the connection between development and inheritance. The Causalist School was defied by the statisticalist reading of natural selection and its emphasis on individual causation.

The second aim was only indirectly presented while challenging the MS. I certainly did not introduce any alternative biological theory, but just some crucial and alternative ideas that arise when questioning the MS pillars. Such ideas constitute new pillars in evolutionary theory. The view behind these new pillars will be central in Part III and my aim to understanding the teleological explanations in development.

New pillars from the Developmental Turn

(i) *Explanatory Internalism:*
Individual organisms are active agents in development. Phenotypic outcomes and organismal activity are not regulated by sub-organismal sources (genes) or designed by supra-organismal entities (evolved populations). The role of organisms in development comes to the for once we acknowledge different advances in contemporary developmental biology: (i) the complexity of genetic activity pictured by the post-genomic era and the need of taking the organism as the proper unit of developmental anlaysis; (ii) the plasticity of developmental pathways and developmental outcomes as the main source of phenotypic variation; (iii) the capacity of organisms of adaptively constructing their niche (both externally and internally) during ontogenesis, and (iv)

the self-organized properties of living systems. The result of these advances is that intrinsic developmental processes are a major adaptive force in evolution.

(ii) *The developmental theory of inheritance:*
The phenomenon of inheritance is not depleted by Replicator Biology. While this input-output view of inheritance works at the populational level concerning measures on heritability, it is silent concerning the processes that produce the cross-generational resemblance. The search for this processes demands an developmental theory of inheritance. Accordingly, inheritance is not the transmission of units of evolved infomration needed to produce adaptations. Rather, inheritance is about the transmission of developmental mechanism than ensure the stability of traits across generations. An developmental theory of inheritance is aligned with the idea that development is not about the unfolding of evolved information but about the recreation of information during a life cycle.

(iii) *Individual causes of evolution:*
Aligned with a statisticalist view of natural selection, evolutionary causes are proximate causes. Evolution is a consequence of organisms struggling for life. The causal process for building adaptive complexity is development, not evolution. In this vein, the DT intends to unify the core ingredients of natural selection at the ontogenetic scale: inherited (Darwinian) fitness variations are caused by ontogenetic processes. Importantly, this does not neglect a central explanatory role for populational explanations; rather this view calls for a distribution of explanatory roles. While proximate analysis is concerned with the causes of adaptive evolution, ultimate, statistical analysis regards how populations change due to heritable variations on (trait) fitness.

Chapter 5

Does ET solve Brentano's Problem?

> Speculation about adaptive significance is a favorite and surely entertaining ploy among evolutionary biologists. But the question, "What is it for?" often diverts attention from the more mundane but often more enlightening issue, "How is it built?"
>
> Stephen Jay Gould, 1983, 152.

In this chapter, I present three challenges to etiological teleosemantics (ET). As expected, they are connected with the three challenges posed to the MS in the previous chapter. The link is quite clear. In Section 3.2.3, I showed how ET rests on MS's pillars. As I challenged these pillars, now it is time to examine the consequences of this challenge for ET. The structure of this chapter follows this argumental strategy. In Section 5.1 I shall focus on the challenges to ET that arise once we abandon Explanatory Externalism. In Section 5.2 I will analyze how the challenges posed to Replicator Biology haunt ET with many troubles concerning its explanatory capacities. Finally, in Section 5.3 I will present a challenge to ET based on a statisticialist view of natural selection.

ET has been severely challenged since its inception. This, in part, explains its improvements and the many efforts that new and old scholars devoted to refining teleosemantics over the years. However, the challenges presented here are, in general, not the classical ones in the literature. So I will opt for a relatively novel way to defy ET. Classical arguments against it concern the Swampman scenario (Davidson, 1987), *disjunctionitis* (Neander, 2017b, 149), content indeterminacy (Fodor, 1990), or the problem of novel contents (Garson & Papineau, 2019). I won't deal with them here. My view on ET is entirely motivated by discussions in contemporary theoretical biology; that is, concerning the *teleo* side of teleose-

mantics.

5.1 Functions beyond externalism

In this section, I will present different challenges to the externalist roots of etiological functions. The idea is quite simple: to the extent that internal, organismal phenomena are also relevant in explaining adaptive complexity, etiological functions based on externalist explanations cannot be fully adequate. Trait functionality cannot purely be the product of a *Panglossian* world (Gould & Lewontin, 1979), i.e. the result of evolutionary design.[1]

Some of the challenges relate to alternative theories of biological functions. To be sure, that of biological functions is a complex issue in the philosophy of biology (cf. Ariew, Cummins, & Perlman, 2002, Krohs & Kroes, 2009, and Garson, 2016, for an overview of the issues). The criticisms concerning evolutionary mismatch (Section 5.1.1) are usually defended from a *modern-history* account (Godfrey-Smith, 1994; Griffiths, 1993). The challenges based on self-organization and global dynamics (Section 5.1.2) gave rise to organizational accounts of functions (Bickhard, 2004; Davies, 2000; McLaughlin, 2000; Mossio, Saborido, & Moreno, 2009). Moreover, the ideas presented in Section 5.1.3 are motivated by Cummins' proposals on functional analysis (Craver, 2007; Cummins, 1975) and an evo-devo account on functions (Amundson, 2000; Amundson & Lauder, 1994; Balari & Lorenzo, 2010; Love, 2007).

In the next subsections, and motivated by Explanatory Internalism (Section 4.1), I will argue that internal, developmental processes are central in adaptive evolution, in such a way that not all functions can be considered to be the result of natural selection. One may retort that etiologists are well aware of the importance of ontogenetic functions, and, as I already pointed out in Section 3.3.1, most etiologists do accept both evolutionary *and* ontogenetic functions. My point, however, is *not* to claim that besides evolutionary functions there are also ontogenetic functions. Rather, I claim that ontogenetic processes are central to evolutionary functions. So it is not about seeing whether ontogenetic functions complement evolutionary functions but realizing that ontogeny is central in adaptive evolution, in such a way that an account of evolutionary functions is not exhausted by externalist explanations.[2]

[1] Chemero (1998) presents a quite similar argument to the one presented in this section. However, while he deals exclusively with the issues of biological spandrels and exaptations, here I attempt to offer many other reasons why internal, organismal processes represent a challenge to the adaptationist roots of ET.

[2] Some scholars prefer to use the label 'devo-evo' instead of 'evo-devo' (e.g. Hall, 2000; G. P. Wagner & Larsson, 2003; cf. Gilbert, 2003 for discussion). If the latter is just under-

5.1.1 Evolution without adaptation

The relevance of internal forces implies that not all traits deserve to be treated from the point of view of an adaptationist logic. This opens the possibility of looking for different evolutionary explanations of phenotypes. In ET, it is posited that the function of a trait is defined by the selection pressure it has helped to overcome. This view deserves revisions insofar as the explanation of the existence and the current functionality of all phenotypes cannot be accountable in such an externalist way. It is, rather, an empirical matter. There may be different evolutionary histories. Not all cognitive traits can be explained on the basis of the Selected Effect Theory of Functions (SETF); this cannot be taken for granted. The explanatory role of organisms in evolution opens the possibility for different mismatches between the SETF and alternative evolutionary histories based on internal causes. Here I analyze some different possibilities in order to see why the evolutionary history of traits is not necessarily the one presented by the SETF.

Spandrels

The amply cited 1979 paper *The spandrels of San Marco* by Stephen Jay Gould and Richard Lewontin (Gould & Lewontin, 1979) represents an inflection point in the history of Explanatory Externalism and a first blow on its underpinnings. The paper soon became a landmark for the critics of adaptationism. Gould and Lewontin's argument touches on different issues around adaptationism. For us, the most relevant one is that internal constraints on development are crucial for explaining the existence of traits and, therefore, externalism cannot be the only source in the explanation of adaptive evolution; other proximate, individual-level ingredients must be part of this explanation. Accordingly, adaptationist thinking —i.e. the idea that an externalist explanation through natural selection of all evolved traits is sufficient— should be rethought.

Gould and Lewontin illustrated the importance of the complexity of developmental processes and their consequence for externalism with a well-known metaphorical example. Imagine an architect designing a church. As it is usually the case, the church has arches, and, necessarily, also columns to support the arches. Consequently, given these architectural features and structural constraints a number of surfaces, known as *spandrels*, will emerge above the point where the extremes of two contiguous arches meet over the supporting column. Spandrels may be used

stood as the study of the evolution of developmental systems, then its main postulates do not necessarily run against the MS, as it in fact the case with the developmental genetics endorsed by the likes of Sean Carroll (Carroll, 2005). The former, however, explicitly assumes that its main goal is to understand the evolutionary consequences of developmental processes. I will henceforth use the expression 'evo-devo' insofar as it is the most common one, but it is important to bear in mind that the central issue at play here is the causal role of development in evolution.

for different functions, such as for painting representation of different religious images. However, spandrels are not part of the architect's intentions, they are not part of the architect's designs, as arches and columns are. Spandrels are structural consequences of the desire of building the church according to the intended design. This example, when we extrapolate it to the biological domain, tries to illustrate why different phenotypic traits could be present in an organism not due to the action of a natural selection process but as a consequence of the role they play during development. In this sense, *biological spandrels*, their presence, cannot be accounted for by natural selection.

The relevance of internal causes also stresses how different traits could be there for other reasons different from the selective ones. As expected, if we look at the cognitive level, we cannot say that all traits deserve selective explanations: many cognitive traits could be part of cognition for other internal reasons. Certain traits could be central in developmental scaffolding; they could operate just during a specific developmental stage, or simply be structurally necessary for the functioning of other traits. The possibility to distinguish between evolutionary selected traits and by-products is certainly even more difficult in cognitive systems due to the complexity and inter-connectivity of cognitive processes.

This challenge can indeed be applied to Sober's distinctions. Recall that, as he argued, the difference between selection-for and selection-of rests on the causal role of the trait function in the process of selection. Previously, I illustrated this distinction by using the example of a salt shaker where there were two kinds of salt: thin and white, and thick and pink. The first one is the only one that can pass through the holes of the shaker. In this example, there was selection-for thin salt because being thin is what makes the causal difference in the process of selection (passing through the holes), while there was selection-of white salt. This example elicits two questions. First, how selection-for is distinguished from selection-of in specific cases? This is an empirical question on the evolution of cognition. It cannot be taken for granted that a certain cognitive capacity was selected-for. I am not saying that evolutionary biology is blind in front of this issues, only that it cannot be solved by stipulation on the basis of a specific definition of biological function. The second question arises when imagining a scenario where the property of being white causes the property of being thin. In this context, whiteness is central in the processes of selection for thinness, even though it is not directly connected to selective pressures. So the functionality of whiteness cannot be accounted for from a strictly externalist position. All in all, the presence of developmental constraints and evolutionary by-products puts the question about the existence of a trait in an empirical domain: not all traits deserve a functionalist explanation of why they are part of nature.

Exaptation and Evolutionary Mismatches

As explained, the SETF defines the function of a trait on the basis of the causal contribution of that function for maximizing fitness during selection processes. Does this mean that the causal contribution of a function during selection processes is the same as the current causal contribution? Not necessarily. Exaptations and Evolutionary Mismatches illustrate why the functions posited by the SETF do not necessarily map into the current functions of a trait.

The phenomenon of exaptation was first described by Stephen Jay Gould and Elizabeth Vrba (Gould & Vrba, 1982). It refers to the fact that, in many cases, the current function (or functions) of a trait are different from the selected function of that trait during natural selection processes; that is, when traits have "evolved for other usages (or for no function at all), and later 'coopted' for their current role" (Gould & Vrba, 1982, 6). This scenario can happen, for instance, when there are harsh environmental conditions, in a way that traits must rapidly accommodate their functioning. Also, it can arise from inner changes in other parts of the organism. If trait X influences trait Y, this may result in modifications on X that bring about a new function for Y. In this case, again, the selected function of Y is not the same one as its newly acquired function. As expected, the main problem for SETF is that the etiological function of a trait —based on selection processes— needs not be the same function as the current one.

A related issue is that of the Evolutionary Mismatch. Evolutionary Mismatches occur when "a trait that evolved in one environment becomes maladaptive in another environment" (Lloyd, 2021, 32). Here again, we have a mismatch between the evolutionary history and the current functioning of a trait. Current maladaptations could be a consequence of changing environmental conditions without natural selection adjusting such novel situations. New, sometimes abrupt —as in the case of species invasion or drastic environmental changes— modifications change the functionality of a certain trait in a way that accommodation is needed to overcome the adverse condition.

In both cases, the difference between selected functions and current functions invites a criticism of the SETF. Principally, the difficulty lies in how selected functions should be empirically established. The point is that we cannot simply apply a naive logic and think about which prehistoric environmental problem gave rise to the current functions of a trait. Ancient environmental conditions need not mirror the current functioning of a trait. Going in the opposite direction, that is applying the technique of 'reverse engineering' (Dennett, 1990) to figure out the evolutionary function just on the basis of the current environmental conditions is not satisfactory either. This kind of logic, paradigmatic of adaptationist thinking in Evolutionary Psychology and Behavioral Ecology (e.g. Barkow, Cosmides, & Tooby, 1992; Pinker, 1997), should be abandoned (Lloyd & Gould, 2017). Current

cognitive functions need not be solutions to Pleistocene problems, nor current environmental problems are a direct path towards the evolution of mind in the Pleistocene. So what exaptations and Evolutionary Mismatches illustrate is the difficulty of establishing the proper function of a trait: empirical studies based on current functioning need not map into the selected functions, and studies about the evolution of traits do not necessarily shed light on the current function of a trait.

Inherency

In Section 4.1.2, I introduced self-organization as an adaptive force in evolution distinct from natural selection. The relevant here point is that self-organization is not exclusive of biological systems. Certain non-living systems are capable of producing emergent patterns by exchanging matter and energy with the environment. Newman (2022b) defines 'physical self-organization' as the property possessed by those systems in which emergent patterns can be explained in terms of the physicochemical properties of their parts. Physical self-organization is produced by 'generic mechanisms' (Newman & Comper, 1990) present in both living and non-living systems, which are governed by physical laws, particularly, by thermodynamics and the principles of dynamical systems theory.

As it is usually assumed, and we will see in Chapter 6, living systems are a distinct kind of self-organized system. Particularly, the creation of self-organized living systems cannot be reduced to generic processes and cannot be accounted for just in terms of physical laws acting on the parts of living systems. The principal difference between living and non-living self-organization is that an agentive and functional dimension emerges in the former. As Kant pointed out when he introduced the concept of self-organization, the distinct purposive nature of living beings is tied to the fact that self-organization is regulated by the system itself in order to stay functional and alive. Non-living systems, on the other hand, do not require any functional language to understand their self-organization; no agentive or functional dimension emerges in nonliving, physical self-organized systems.

The demarcation between living and non-living self-organized systems requires much attention. This demarcation will be explained in detail in Chapter 6. But once I stressed the physical nature of some self-organized systems in nature, I am able to articulate the central claim in this section: that different traits emerge during evolution and development as the result of generic mechanisms acting on the physicochemical bases of cells; i.e. many traits may be originated due to physical reasons only.

This idea has been developed by Stuart Newman under the label of *Inherency* (Newman, 2021) and it finds support in many empirical studies (Forgacs & Newman, 2005; Newman, 2012, 2022b; Newman & Comper, 1990; Newman, Forgacs,

& Müller, 2003; Newman, Glimm, & Bhat, 2018). According to Newman's definition, "inherency means that certain structural motifs (e.g. tissue layers, lumens, segments, appendages) can be readily generated by physical organizing forces acting on tissues masses" (Newman, 2021, 121). His idea, therefore, is that generic physical mechanisms are central to the evolutionary origin and development of many traits. My point here, following Newman, is that the production of traits due to generic mechanisms is independent of any externalist or functional description. Physical self-organized (anatomical) patterns emerge due to the physicochemical composition of their parts, they are not the result of selection processes but of the intrinsic dynamics of developmental systems: "if morphological novelties arose by means other than cycles of gradual change, evolution of form cannot mainly be a question of fitness and relative advantage, but rather of development and its transformations" (Newman, 2022a, 199). In other words, the evolution of many traits cannot be just reduced to the understanding of their causal contribution to fitness maximization during selection processes (as etiologists claim); not all traits deserve an etiological explanation. The physicochemical constitution of living beings may provide an alternative non-etiological answer to the presence of a trait in nature, insofar as "major pathways of evolution are determined by physical law, or more specifically by the self-organizing properties of biomatter, rather than natural selection" (Edelmann & Denton, 2006, 578-580).

Note that this situation is similar to the case of biological spandrels. Spandrels are part of developing systems not because of their contribution to overcoming external pressures but for intrinsic, formal reasons. Also, the origin of self-organized patterns during evolution is not within the scope of Explanatory Externalism but of Explanatory Internalism: "Self-organized order is spontaneous pattern from within; the order of selection is additive order from without" (Edelmann & Denton, 2006, 588). This is the reason why Newman concludes that "inherency is not merely complementary to the Darwinian paradigm, but is at odds with it" (Newman, 2021, 130).

5.1.2 Functionalism vs structuralism: a revival

Some approaches within developmental biology, particularly Evolutionary Developmental Biology, have criticized the biological framework underpinning ET. The etiological determination of natural kinds does not fit well with the evo-devo picture of evolution rooted in a structuralist biology. In particular, two interconnected issues arise from an evo-devo viewpoint: developmental homologies as an alternative account of trait taxonomy and the structural/formal constraints on organisms as the proper units of functional attributions.

Developmental homology

First, how traits should be classified? How can we say that two traits belong to the *same* trait type? This issue connects with the old dispute between Étienne Geoffroy Saint-Hilaire and Georges Cuvier (Appel, 1987), and later structuralist-functionalist confrontations. Richard Owen was responsible for developing a homological theory of traits based, among other criteria, on morphogenesis. This view has been recently revived by a number of scholars (Brigandt, 2002; Brigandt & Griffiths, 2007; Love, 2008; G. P. Wagner, 2001, 2014). The main lesson is that, following the structuralist foundations of evo-devo, the notion of (developmental) homology should be taken as the proper unit of trait classification, instead of the functionalist approach based on Darwinism. But first, let's introduce the taxonomy of traits that arises from an evolutionary, historical view of homology as it was defended by Darwin, to contrast it then with the alternative, 'Owenian', developmental account of homology.

Darwinian taxonomies are based on evolutionary histories. The Unity of Type —what determines when a trait is the same one in different members of the same species and between species— is accounted for by common descent. How can we say that two traits belong to the *same* trait type? Sameness is the result of shared phylogenetic histories. Two traits in different species are the same if this trait is also present in the most recent common ancestor of both species. Following DiFrisco (2021, 1), "two characters in distinct organisms or taxa are homologous if they are genealogically connected by continuous descent from a common ancestor that had the same character". This is the historical definition of homology. The explanation of trait classification relies on the processes of natural selection based on descent with modification. The functionalist roots of Darwinian taxonomies lie in the notion of adaptation. As noted (cf. Sober's quote on page 47), an adaptation is connected with the function that a trait type performed during evolution by natural selection. Explaining the Unity of Type by common descent is tied to an identification of traits on the basis of the function that these traits performed during evolution. Shortly, sameness is the result of the common descent of evolutionary adaptations.

The historical view of homology was a reinterpretation of the ahistorical account proposed by Richard Owen, and recently retaken by different evo-devoists, especially by Günter Wagner (G. P. Wagner, 2001, 2014, 2016). From this viewpoint, structure is prior to function in the classification of homologous (and serially homologous) traits. This view is ahistorical insofar as it concerns the developmental processes involved in morphogenesis —i.e. the development of form. Rather than providing a historical view of functions, evo-devoists supply it with a taxonomy based on an ahistorical view of the development of form. What determines homology is therefore the presence of the same developmental processes in differ-

ent organisms of the same or different species. In a nutshell, ahistorical homology is based on morphogenesis, an individual-level phenomenon, while historical homology is based on tree-thinking (G. P. Wagner, 2016), a cladistic, phylogenetic account of traits. In one case, morphogenesis —as a developmental process— explains Unity of Type, while in the other case, phylogenesis explains Unity of Type. The central differences are three: evo-devo homologies are ahistorical, developmental, and based on the construction of an organism's structure. Darwinian and Neo-Darwinian homologies are historical, evolutionary, and determined by the function of traits.

Evolved by structural motifs... and then function

The second issue, connected with the previous one, is that functionality is always tied to the structure of the system (cf. Bock and von Wahlert (1965) for a seminal work). This issue is linked with the traditional debates between functional and structural biological theories. Does function guide the origin of structure or the other way around? As explained, the structuralist underpinnings of evo-devo search for the origins of a trait in the very process of morphogenesis (cf. Amundson, 2005, for a clear exposition of Darwinian functionalism vs. evo-devo structuralism). In doing so, as in the case of spandrels and inherencies, the presence of some traits is revealed as the result of structural constraints during development; the function of a trait during selection does not produce new structure, but the origin of organic structures constraints and allows different functions (Amundson, 1994). Instead of waiting for an opportunistic random variation that functions properly during gradual processes of cumulative selection, "most physiological functions appeared in animal lineages abruptly and essentially ready-made [by the origin of structural variations and novelties]" (Newman, 2022a, 199). The structure of a trait predates its functionality: "Inherency [as spandrels] makes generation of form ontologically prior to its uses" (Newman, 2021, 122).

As expected, this perspective suggests a view of function alternative to the SETF. Evo-devoists see trait functionality as the result of the structural and organizational properties of living systems. The main idea at play is that a trait can exhibit different functionalities as a consequence of its formal properties. The unit of analysis is not the function, but the organic structure that enables different functions in different contexts and connections with the rest of the system. For example, Love (2007) argues that the functionality of a trait must be accounted for in terms of the activities that a certain structure allows. He thus defines *activity-functions* in contrast with *use-functions*, which provide a valuative dimension for a particular functionality. Balari and Lorenzo (2010) coined the term *functional-ability* to refer to the repertoire of possible functions a trait may have due to its formal properties. In these cases, the organic structure is prior to proper func-

tioning. First, in definitional terms: sameness is based on morphological rather than on evolutionary/functional terms. And second, the usefulness of traits is a consequence of their formal properties. The challenge to etiological theories is quite clear. First, etiological functions are based on a functionalist view of trait characterization. Secondly, etiology neglects the role of inner constraints in the attribution of functions. Finally, etiology does not posit that the activity of a system is the consequence of its structural properties, but a consequence of evolutionary histories.

Certainly, in many cases, an evo-devo position is connected with an eliminativist view of teleofunctions (but not always, e.g. Newman (2022a)). From this viewpoint, there are no functions that traits must perform; there are no *proper functions*. There is no normative valuation, nor purposefulness in the activities of organisms (e.g. Amundson, 2000; Amundson & Lauder, 1994; Balari & Lorenzo, 2010). That is why Cummins' stance is usually adopted in structuralist views. As already pointed out, Cummins-functions are not teleological, because they are defined only on the basis of the activity that a trait performs in the context of the system. As was defended also by Searle (1995), teleological attributions are extrinsic, not proper to the system but based on an external observer. Under this structuralist position, teleofunctions are not part of nature. They might be valuable epistemological tools, but the activity of a trait is not connected with any purpose or norm but just with the formal constraints of the system.

I accept most of this evo-devo insights. I submit that externalist attributions of functions are misguided. There are two basic sources for externalism: natural selection and the observer. The first one was already criticized. The second one, defended by Searle (1995), is a relativism of sorts: the function of a trait would be different according to the aims of the scientists and the context of inquiry. I agree that none of these externalist positions allows for a suitable treatment of teleofunctions. However, if I aim to explain how teleofunctions can find their place within the DT, and evo-devo is one of the main frameworks within the DT, then I need to accommodate teleofunctions within the structuralist picture of evo-devo. In other words, there are functions beyond external stipulation; i.e. there are intrinsic teleofunctions. I will argue in Part III that a structuralist view can be reconciled with an intrinsic account of teleofunctions.

5.1.3 Indeterminacy and Natural Selection

Let's go back to Gould and Lewontin's idea of spandrel in order to call the attention to a different problem, presented in Fodor and Piattelli-Palmarini (2010), and connected to the Content Determinacy Challenges (Fodor, 1990). If all hearts make noise when they pump blood, pumping blood entails noise-making. It goes without saying that it is pumping blood what mainly contributes to fitness and

what makes hearts to be present in nature, not the noise they produce, even though noise, as a spandrel, may be of some use (for pulse-detection, for instance). But the issue is not whether we can discern between such co-extensive functions.[3] The point is whether natural selection can. Let's present the situation with our example of the salt shaker with two kinds of traits: thin and white salt, and thick and pink salt. The 'fitter' trait is the thin and white one because being thin allows the salt to pass through the holes of the shaker. In this example, the populational sample is the salt in the shaker and the salt that is white is also thin and vice versa. Consequently, the properties of thinness and whiteness are co-extensive. Remember that this example was introduced to present Sober's distinction between selected-for effects (thinness) and selected-of effects (whiteness). But the question is, can natural selection distinguish between co-extensive functions in such a way that it selects-for one of them while selecting-of the other one?

It is not so difficult to show that it cannot. The main point is that co-extensive functions cannot be distinguished in selection processes insofar as they do not provide any variation on fitness. Selection just preserves those traits that contribute to the organism's increment of (inclusive) fitness (i.e. leave more offspring). But it does so without looking at how such a trait maximizes fitness. This is usually referred to as the blindness of natural selection: natural selection does not take into account the causal mechanisms that make a trait an important contribution to fitness maximization, rather natural selection just 'sees' whether fitness increases or not due to having a particular trait. If, as in the example of the salt shaker, the trait has two 'functions', thinness and whiteness, natural selection is not capable to tell which of them is doing the causal work. It just perpetuates the fitter trait.

Clearly, *we* can distinguishes between selection-for and selection-of. For instance, we can analyze how being thin makes thin and white salt be selected. Being thin is a causal factor that contributes to passing through the holes of the shaker. But, as explained, natural selection is not capable of providing such kind of analysis. Another strategy is to imagine a counterfactual scenario where there is thin salt that is not white, such that not being white does not affect the fitness of thin salt. But this is tricky. Counterfactual scenarios are imaginable only by intentional systems. Intentionality is needed to discern between co-extensive properties. If it were not so, natural selection would be an intentional mechanism. This position is not easy to defend, insofar as part of the successfulness of ET rests on the non-intentional character of natural selection to naturalize intentionality in cognitive systems. The Darwinian moral was that there is design in nature but without a designer.

[3] I use co-extension not in a strict, logical-mathematical way, but just to refer to the connection between two properties of a trait when the existence of one of them entails the existence of the other.

To illustrate my argument, let's take a look at this lost Platonic dialogue discovered by Fodor (2008, 133) and set in Table 5.1 on the next page.

In Plato's so far lost dialog (left), Socrates argues that if content is determined by intension, we cannot discern co-extensive situations without supposing prior intentionality. Beyond Plato's worries, I believe that the dialog can be also useful in shedding some light on my argument by introducing some minor changes (right). In my argument, when I use the expression *selectionally* distinct I am referring to whether the concept was selected-for or selected-of. The argument runs by noting that traits are selected from their instances —i.e. by cumulative generations. However, due to their co-extensiveness, in each instance, we have both C and C*. So there is no difference between selecting C and C*; therefore, C and C* are not selectionally distinct: whether C is selected-for and C* selected-of (or vice versa) is not something natural selection can determine. The moral is that *co-extensive phenomena are not variations*; the blindness of natural selection makes it incapable to see how a trait causally contributes to fitness maximization. In these co-extension scenarios, the distinction between selection-for and selection-of, which is central in etiology, cannot be the result of selection processes.

5.2 The Phylogeny Fallacy and Replicator Biology

5.2.1 The Phylogeny Fallacy

The Phylogeny Fallacy is the conflation of evolutionary explanations with ontogenetic explanations. Specifically, one commits the Philogeny Fallacy when evolutionary explanations about populational processes are invoked to explain phenomena at the individual, ontogenetic level; that is, phenomena that fall beyond their explanatory scope. An explicit endorsement of the fallacy, and probably the first one, could be Ernst Haeckel's statement that "the theory of descent alone can explain the developmental history of organisms" (Haeckel, 1866, 7). Another way to unveil this fallacy is by looking at Mayr's classification of causes and explanations (cf. Section 2.2.3). In this context, the conflation consists of the use of ultimate causes to explain proximate causes, of ultimate explanations to deal with proximate explanations, or to provide why-answers to how-questions.

The Phylogeny Fallacy was thus baptized by Lickliter and Berry (1990). It is nowadays often denounced by Developmental Systems Theorists (Griffiths, 2002; Oyama, 2000b; Oyama et al., 2001), Developmental Psychobiologists (Gottlieb, 1997; Michel & Moore, 1995) and Eco-Devoists (Lewontin, 2000). The strongest criticisms came however from the American ethological school, which put a strong emphasis on embryological studies of behavioral development. This school chal-

Fodor's discovery:	**My version:**
Q. Do you think that two concepts could be intentionally distinct but coextensive? A. Yes, Socrates. Q. Such that someone might learn one of the concepts without learning the other? A. Yes, Socrates. Q. And such that the concepts in question might both be learned from their instances? A. Yes, Socrates. Q. Very well. Now consider the coextensive but distinct concepts C and C*. Do you not agree that, since these concepts are coextensive, everything that's an instance of C is likewise an instance of C*? A. Yes, Socrates. Q. And vice versa? A. Yes, Socrates. Q. Now tell me: if everything that is C is C* and vice versa, what determines whether it is C or C* that one learns from one's experience? A. Yes, Socrates.	Q: Do you think that two traits could be *selectionally* distinct but co-extensive? A: Yes, Socrates. Q: Such that natural selection might select one trait without selecting the other? A: Yes, Socrates. Q: And such that the trait in question might be both selected from their instances? A: Yes, Socrates. Q: Very well: Now consider the coextensive but selectionally distinct traits C and C*. Do you not agree that, since these traits are coextensive, everything that's an instance of C is likewise an instance of C*? A: Yes, Socrates. Q: And vice versa? A: Yes, Socrates. Q: Now tell me: if everything that C is C* and vice versa, what determines whether it is C or C* that natural selection selected during evolution? A: Yes, Socrates.

Table 5.1: A lost Platonic dialog discovered by Jerry Fodor in 2008

lenged and criticized the idea of instinct already in the early 20th century, through its principal and most eloquent figure, Zing Yang Kuo, who set the scene for the later work of Theodore Schneirla and Daniel Lehrman (Lehrman, 1953, 1970; Schneirla, 1966), and, more recently, of the late Gilbert Gottlieb (Gottlieb, 1991, 2007).

The main locus —but not the only one (cf. Section 5.2.3)— of the Phylogeny Fallacy is *Dichotomic Thinking*, the principal manifestation of which is made apparent in those dichotomic terms that are typically associated with the nature-nurture debate, such as innate-learned, inherited-environmental, or biological-cultural. Different scholars have stressed the unwarranted character of views based on Dichotomic Thinking and dichotomic categories. I will focus here on the three main criticisms that have been wielded against it: explanatory vacuity, empirical inadequacy, and semantic clutter.

Explanatory vacuity

Explanatory vacuity is intimately connected with the Phylogeny Fallacy, a point already stressed by Kuo in his work and in his attacks on the theory of instinct. The most popular theory of instinct is the one proposed by Konrad Lorenz, who distinguished between instinctive behavior and acquired behavior. The distinction rests on two different sorts of learning processes: evolutionary learning processes, and ontogenetic learning processes. In the case of evolutionary learning processes, natural selection biases behavioral variations toward adaptive traits. Species 'learn' about what behavior is adaptive thanks to natural selection picking up the behaviors that maximize fitness. The other source concerns ontogenetic learning. Ontogenetic learning consists of different processes of organismal regulation involving behavioral feedback coupled with the environment. Moreover, these two learning processes provide different kinds of information for development: phylogenetic information and ontogenetic information. The information achieved through evolutionary process is labeled *phylogenetic information* while *ontogenetic information* arises via ontogenetic learning. The distinction between instinctive and acquired behavior rests on these dichotomies. Instinctive behavior develops by maturation through a rigid context-free process based on phylogenetic information acquired through *phylogenetic learning*. Acquired behavior, on the other hand, is context-sensitive to the ontogenetic information that is acquired during ontogenetic learning. What is important to note here is that information in development comes from two different sources: the evolutionary and the ontogenetic ones. This is the core of Dichotomic Thinking.

Kuo's view on instinct could be perceived as an anachronism, especially if we take into account that most of his work was developed well before Lorenz's theory. However, Kuo attacked the idea of instincts as obtained through a process of

maturation independent of environmental inputs and accounted for in terms of evolutionary explanations, the view defended later on by Lorenz (and also by Dennett (1995); Skinner (1953, 1957), for instance). Only later, Daniel Lehrman (Lehrman, 1953), acknowledging Kuo's influence, was the direct sparring of Lorenz's theory. So, what lessons can we derive from Kuo's critique?

The main lesson is that instinct theories lack in genuine developmental explanations:[4] "The use of the distinction generates in researchers the false illusion that certain important empirical questions have already been answered" (Bateson & Gluckman, 2011, 129). Or, as Oyama (2000b, 159) pointed out, "[i]t feels right, but it doesn't explain anything". A genuine developmental explanation must explain how traits arise through different ontogenetic stages. This explanatory aim is achieved by looking at the mechanisms of development; that is, at how parts interact, at how different causes act and guide development, and at the many sources of information and the complexity of developmental dynamics. Instinct theory, on the other hand, avoids looking into the mechanism of development (it black-boxes it). In a nutshell, claiming that a trait develops by phylogenetic information does not inform about the mechanisms of development. As Kuo eloquently claimed, instinct theory, and the dichotomic view in general, is a "finished psychology" (Kuo, 1922, 345). Theories of instinct provide no information about how behavioral traits develop during ontogeny; therefore, "to call an acquired trend of action an instinct is simply to confess our ignorance of the history of its development" (Kuo, 1921, 650).

I shall use the phrase *Kuo's Lesson* to refer to the conflation of explanations that gives rise to the Phylogeny Fallacy. Paul Griffiths coined the expression 'Lehrman's Dictum', which boils down to essentially the same idea:

> The idea of genetic information [including Lorenz's phylogenetic information], like the idea of innateness, is a Trojan horse that helps to disguise an evolutionary explanation as a developmental explanation, and obscures the fact that no actual explanation of development has been produced. (Griffiths, 2013, 23)

[4] Another relevant insight is that many non-obvious resources can play an important role in development. Sometimes instincts are established from isolation experiments: the result of the same developmental outcome even in the absence of a particular environmental input. This stance is mistaken for two reasons. First, isolation experiments deal with a specific and expected developmental cause, but isolating the organism from a developmental cause is not the same as isolating the organism from its environment (this is definitively impossible). Secondly, the lack of a particular environmental input could be accounted for by the plasticity of developmental processes and their capacity of producing robust outcomes (Bateson & Gluckman, 2011).

Empirical inadequacy

What information is given about a developmental process when a trait is said to be innate? This criticism concerns the empirical inadequacy of most attempts to answer this question. For us, it will be easy to appreciate this point insofar as most of the proposals are anchored in a MS framework.

A classical answer to the question opening this section is that innate traits are genetically programmed. From a post-genomic viewpoint, we know why this is problematic. First, there are no such thing as a genetic code for traits. The G×P map is contingent upon the environment, as the different norms of reaction set in Figure 4.1 on page 92 illustrate. Second, and crucially, all traits depend both on genetic and epigenetic causes. As Mary Jane West-Eberhard once wrote, genes, without the cellular machinery, "are among the most impotent and useless materials imaginable" (West-Eberhard, 2003, 93). Genes are central in the production and reproduction of any cell, as well as in the process of cell differentiation. So using a genetic parameter for defining innateness overlooks the fact all traits depend on a myriad of developmental causes.

However, another reading of genetic programming may possible. It could be understood as stating that genes are a specific kind of cause in development: genes provide the information needed for developing particular traits. So even if there are many other resources in development, genes are the only ones that provide the relevant information for developing innate traits. However, once again, this view stands in contradistinction with the postulates of post-genomics. In essence, the central claim of post-genomics is that non-genetic resources are also crucial sources of information in development. Specifically, understating genetic activity as a subordinate of the cellular context is not just saying that genes need the material basis of the cell, but also that the cell provides the specifications about which outcomes would be produced. So the role of being an informational source of development is not exhausted at the genetic level. All in all, the genetic parameter for defining innateness is classically rooted in a misguided view of development based on the causal powers of genes specifying developmental outcomes. It might work with the molecular gene, but not with reactive genomics.

Alternatively, one could claim that innate traits depend on inherited systems, in opposition to ontogenetic information specific to each lifespan. The problems here are two. First, these proposals usually take inherited information to be exclusively genetic; but, as already noted in Section 4.2.1, inherited information is not uniquely genetic. So by tying inheritance to genetics, extended inheritance is ignored, and besides it brings about the aforementioned problem concerning the genetic definition of innateness: there are no such things as genetic inherited traits. The second problem is that, even if we assume an extended view of inheritance, we would also run into trouble. Nativist talk usually comes together with

the attempt to separate inherited causes of development from environmental ones, inner causes form external ones. However, as soon as we realize that exogenous resources can also be inherited, this enterprise loses much of its sense, because the dichotomy between inherited and non-inherited terms does not clearly map into the distinction between inner and external causes of development. A further problem of defining innateness in terms of extended inheritance will be discussed in Section 5.2.3 in relation to Shea's theory of inherited representations.

Another empirically mistaken proposal is to say that a trait is innate if it develops in different environmental contexts, which typically elicits the conclusion that its developmental trajectory is independent of external resources. This kind of proposal is usually related to isolation experiments: to isolate the developing system from a particular environmental resource and see what the outcome is. If the outcome remains unchanged, then such an environmental resource is irrelevant in development; if different environmental contexts do not produce variation, we have an innate trait that does not depend on the environmental context. I think that there are two problems here. Firstly, as it was already noted by Kuo in his experiments with chick embryos (Kuo, 1932), there might be non-obvious epigenetic resources that participate in development. As developing organisms cannot be isolated from their environment, but just of some environmental resources, there is the possibility of other environmental resources being involved in development. Secondly, and most importantly, to obtain the same outcome in the absence of an environmental resource does not necessarily mean that this resource is not playing a role in development. This is well recognized if we acknowledge the robustness of outcomes as the result of the plasticity of developmental pathways. A developmental resource may be central in development but, if this resource is not present, the developing system may nonetheless find an alternative path towards the same outcome. A robust outcome in front of a changing environmental context does not entail that the trait is developed 'from within', but maybe that the developing system can compensate for different environmental scenarios.

In general, from the view of an individual's lifespan, all developmental resources are in an equal position —an idea labeled as the *Parity Thesis* by Griffiths and Gray (1994). This is the reason why the distinctions behind Dichotomic Thinking are not based on the analysis of developmental processes but on a distinction between evolutionary and ontogenetic causes. Genuine developmental explanations are about the mechanisms of development, not about the classification of traits on the basis of their populational/evolutionary properties.

Semantic clutter

Semantic clutter concerns the difficult task of providing clear and useful definitions for dichotomic terms. For example, what does innate mean? Well, it depends.

It depends on the time, the context and a myriad of other factors, as revealed by the abundant literature on the topic (e.g. Bateson & Mameli, 2007; Griffiths, 2002; Lorenzo & Longa, 2018; Mameli, 2007, 2008; Mameli & Bateson, 2006, 2011; Wimsatt, 1986). As Wittgenstein (2010) famously defended, the meaning of some expression is determined by a cluster of interrelated concepts, not by an exact definition or a definite description. But is innate a cluster concept or it rather is more like a *clutter* concept? The critical point is that many of the properties that are typically associated with innateness are not necessarily observed together in the living world; that is, sometimes we find one of these properties but not the others. Some of the options are: being an adaptation, present at birth, being genetically inherited, being robust or canalized, or being genetically based. The problem is that these properties do not come in a package. There are many cases where they appear dissociated. While this does not necessarily mean that *innate* is meaningless, it does nonetheless suggest that there does not seem to be a clear cluster of interrelated concepts that would allow for a scientific account of innateness. Consequently, why should we appeal to the general property of being *innate* when we can refer to each property separately? As Griffiths put it:

> If a trait is found in all healthy individuals or is pancultural, then say so. If it has an adaptive-historical explanation, then say that. If it is developmentally canalized with respect to some set of inputs or is generatively entrenched, then say that it is. If the best explanation of a certain trait differences in a certain population is genetic, then call this a genetic difference. If you mean that the trait is present early in development, what could be simpler than to say so? If, finally, you want to 'blackbox' the development of a trait for the purposes of your current investigation then saying so will prevent your less methodologically reflective colleagues from supposing that you think the trait is [...] innate. (Griffiths, 2002, 82)

The moral is that dichotomic categories conform a clutter of properties that lack a clear definition. And here is where the connection with empirical adequacy is made explicit: if genes really played the role of orchestrating development, perhaps a genetic definition of innate trait would work, but the complexity of development and the many sources of inheritance prevent the clutter from becoming a useful cluster.

To conclude, let's note that Dichotomic Thinking would not be safe just by assuming an interactivist picture based on evolutionary and ontogenetic explanations. Again, there are two kinds of interactivism. *Type 1 Interactivism*, or the so-called 'interactionist consensus', is the idea that all traits result from the interaction of innate —evolutionary— and learned —ontogenetic— resources. Type 1

Interactivism also promotes a misguided view of development based on two different kinds (ultimate and proximate) of developmental causes: "This 'interactionist consensus', however, perpetuates the nature–nurture debate by maintaining its inherent dichotomy" (Stotz, 2008, 360). Unlike the former, *Type 2 Interactivism* is free from the Phylogeny Fallacy and related problems. According to this view, phenotypic outcomes are the result of the interaction of many developmental resources at the individual level —genetic, cellular, extracellular, and exogenous. In this kind of interactivism, there is no plurality of explanatory strategies but a plurality of proximate developmental resources (cf. Oyama, 2000b, for a detailed criticism of Type 1 Interactivism and a defense of Type 2 Interactivism).

5.2.2 Dichotomic Thinking in etiological teleosemantics

In this chapter I will present two arguments why in my opinion ET endorses the Phylogeny Fallacy. The first one, to be developed in this section, is an argument by ostension, where I just pinpoint those places where ET appeals to Dichotomic Thinking.[5] Therefore, the strength of my argument is relative to the scope of my analysis. I will deal with the classic characters of teleosemantics: Millikan, Papineau, Neander, Shea, and Dretske. The second argument, to be presented in Section 5.2.3, is independent of any particular proposal in ET; i.e. it is a problem inherent to ET itself.

Dichotomic Thinking takes two forms in the teleosemantic literature:

1. The explicit use of dichotomies.

2. The distinction between evolutionary and ontogenetic functions.

I already explained why dichotomic categories entail the Phylogeny Fallacy. As for Item 2 above, the problem lies in appealing to ontogenetic and phylogenetic functions to explain representational capacities. The problem is not the demarcation itself but the idea that they can be joined in the explanation of representational capacities. The result of this view is a mix of explanatory strategies and causal sources. We obtain some representations from evolutionary processes while others from ontogenetic ones. This is Dichotomic Thinking. Biological, innate, or inherited representations are a consequence of evolutionary (etiological) functions, whereas cultural, learned, and environmentally-specific representations are accounted for in terms of ontogenetic functions.

So my argument by ostension essentially boils down to the fact that the literature on ET typically invokes Item 1 and Item 2. Concerning Item 2, I already

[5] I took the idea of *argument by ostension* from Neander (2017b).

noted in Section 3.3.1 how ET systematically appeals to different sources for functions. So, regarding multiple functions, my argument for ostension refers to that section.

As for Item 1, the use of dichotomic categories in explaining representational capacities, there are different periods in the history of teleosemantics in which such categories have been appealed to. In Table 5.2 on the next page I compiled some illustrative expressions used in teleosemantics typically associated with dichotomic thought.[6]

Moreover, the use dichotomic terms has a direct connection with the appeal to multiple functions. In other words, Item 1 and Item 2 are interconnected. As Table 5.3 on the facing page illustrates, Dichotomic Thinking in teleosemantics results into *two groups* of interrelated conceptual classifications. On the one hand, the first is explained in terms of the SETF; here populational explanations involving natural selection are invoked to account for the construction of innate representations: "In the case of innate representational capacities, the relevant selection process is neo-Darwinian natural selection, so that the function of a system is to do whatever ancestral systems did which caused systems of that type to be preserved and/or proliferated in the population" (Neander, 2006b, 381). On the other hand, the second is explained in terms of ontogenetic functions, where individual-level phenomena involve learning or persistence to construct acquired representations.

Importantly, the relation between both groups may be characterized as a form of Type 1 Interactivism. For instance, Millikan claims that "[i]nner states, such as the perceptual and cognitive states of organisms, can have proper functions that vary as a function of environmental input to the genetically programmed systems responsible for producing them. Unlearned behaviors can have proper functions that are either variant or invariant with respect to environmental input" (Millikan, 2000a, 86). Even more explicitly, Dretske states that "the old nature-nurture dichotomy is too simple. Behavior is the product of a dynamic interaction between genetic and environmental influences. The innate and instinctive is inextricably intertwined with the learned and the acquired" (Dretske, 1988, 31).

This variety of interactivism entails is that both groups causally contribute to the development of traits, not necessarily one by one, but in relation with each other: selected functions may be complemented by ontogenetic functions, innate capacities may be fine-tuned by acquired capacities, innate information interacts with acquired information to produce complex representations, and inherited material is intervened by the environmental context of development. In sum, the

[6] Additional examples can be found in Millikan (2006), Millikan (2017, ch. 5), Millikan (2002), Garson and Papineau (2019), Papineau (2017, 118), Neander (2007, 550, 559, 560), or Dretske (1988, 47).

Dichotomic expressions	References
Innate perceptual–cognitive mechanisms	Millikan (2006, 109)
Genetically programmed systems	Millikan (2000a, 86)
Innate skills, abilities	Millikan (2000b, 54. 63, 65)
Innate, hard-wired belief-forming abilities	Papineau (1984, 557)
Innate capacities	Neander (2017b, 82, 101)
Innate sensory-perceptual systems	Neander (2017b, 166)
Innate information, representations	Neander (1995b, 111-112)
Innate behavior	Dretske (1988, 123)
Behavior causally explained by the genes inherited	Dretske (1988, 92)
Genes coding for behaviour	Dretske (1988, 123, 125)
Rigidly programmed behaviour	Dretske (1988, 125)

Table 5.2: A (non-exhaustive) list of dichotomic expressions found in the literature on teleosemantics.

Populational-level SETF	Individual-level Ontogenetic functions
Natural Selection	Persistence, Learning
Innate representations, information, abilities, capacities	Learned representations, information, abilities, capacities
Genetic or cultural inheritance	Environmentally induced
Non-intentional (Dretske)	Intentional (Dretske)

Table 5.3: Dichotomic Groups in Teleosemantics. On the populational-level side, SETF constituted by natural selection processes construct innate representations based on genetic or cultural inheritance. On the individual-level side, ontogenic functions based on persistence or learning construct learned representations from environmental inputs throughout development.

populational, evolutionary underpinnings of development interact whit the individual idiosyncrasies of development. As already explained, Type 1 Interactivism is problematic. The trouble is not about distinguishing different developmental causes. Rather, the problem is that dichotomies are grounded in different sorts of causes (ontogenetic and phylogenetic) and explained by different sorts of processes (natural selection and learning).

Many of Dretske's insights will be re-examined in Part III. He supported an informational teleosemantics based on individual-level processes. He also attributed some sort of representational capacities to non-cognitive systems (such as plants) by noting the importance of signals (or indicators) in plants' lifespan, for example. However, as explained, he traced a barrier between intentional and non-intentional systems. This barrier is based on Dichotomic Thinking: if the representational systems result from evolutionary processes, then no intentionality is posited; but, if learning takes place, goal-directedness arises. Like Dretske, I will also argue for an informational teleosemantics based on individual-level processes, and I will stress the importance of signals and indicators in development. However, I will need a different criterion to define the barrier.

As noted, Shea's view on etiological function is complex and relevant for my discussion. He defends Dichotomic Thinking but he believes that classical concepts such as innate, instincts, or genes coding for traits are misguided terms. Rather he supports that his theory of *inherited representations* is suitable for anchoring Dichotomic Thinking and accounting for "innateness-related properties" (Shea, 2012a, 2012b). Moreover, his proposal also rejects many ideas traditionally associated with the MS by stressing the importance of non-genetic developmental resources and non-genetic systems of inheritance (Shea, 2007b, 2013). His view will be discussed in next section.

5.2.3 The Phylogeny Fallacy and Replicator Biology

To complement the argument by ostension, in this subsection I will argue that ET finds itself in a difficult situation for it being rooted in Replicator Biology. Such a situation generates a tension between two options: either it endorses the Phylogeny Fallacy (a misguided *explanans*) or it abandons its aim of accounting for the intentional gap (i.e. gives up its central *explanandum*). I will begin by tracking down first the connection between Replicator Biology and the Phylogeny Fallacy and, second, the connection between Replicator Biology and ET. Next I will introduce a central *desideratum* for any teleosemantic project, the *Actuality Desideratum*. Based on these ideas, I will argue that ET needs to change its *explanans* or its *explanandum*. I will close with an analysis of Shea's proposals on inherited representations and why it is affected by the same flaws.

The link between Replicator Biology and ET was already described in Sec-

tion 3.2.3. Replicator Biology provides the necessary theory of inheritance for the MS solution to Kant's Puzzle on which etiology rests. Etiology has to do with the adaptive advantage of traits fitness in a population and with how natural selection operates on these traits during generations thanks to reliable inherited systems. Replicator Biology provides the necessary connection between what is inherited and what is selected. Moreover, it does so without looking into the mechanism of development. Adaptations and etiological functions arise by biasing populations in an adaptive way. Selection is the external adaptive bias. Inheritance secures stability and fidelity and, given the assumption of the Weismannian framework, it is not necessary to understand the causal processes acting between the inherited units and the selection of phenotypic outcomes. This is enough for evolution, this is enough for etiology. Individual-level phenomena are explanatorily vacuous.

As for the connection between Replicator Biology and the Phylogeny Fallacy, it is best appreciated if we take a look at the history of biology. Replicator thinking and Dichotomic Thinking emerged more or less simultaneously and were grounded in the same biological framework. This connection is clearly and accurately dug up by Evelyn Fox Keller in her wonderful *The Mirage of a Space between Nature and Nurture* (Keller, 2010).

The core idea is that both kinds of thinking (replicator and dichotomic) detach inheritance from development. Studies about the evolution of populations do not inform about developmental processes. As Keller notes, this conception involves two relevant conceptual shifts advanced by the MS: from individuals to populations, and from trait to trait differences. While the connection between inheritance and development rests on investigations about *how individual phenotypes* are constructed in each generation, the disconnection between inheritance and development concerns *how many* traits resemble each other in a *population* —which *trait differences* are preserved and which ones are not. As can be noted, and Keller argued, these changes promoted an explanatory shift concerning the questions they answer. While analyzing traits at the individual level answers how-questions, analyzing trait differences at the populational level answers how-many questions. The classical problem of Dichotomic Thinking, which I am blaming ET for endorsing, is to presume that how-many answers are also how answers. One cannot analyze individual traits by just looking into trait differences within populations.

The connection between Replicator Thinking and Dichotomic Thinking is that both avoid getting into the mechanisms of development. Replicator Biology is not a developmental theory of inheritance and, accordingly, it has nothing to say about how development unfolds; it just looks at the connection between inherited units and selected traits. Dichotomic Thinking identifies the role of a resource in development by looking at the phenotypic differences in a population that it brings

about. But it does so without actually looking into the causal processes in which such a resource participates. Both kinds of thinking answer how-many questions, none of them tackles how-questions.

To appreciate the difficult situation that teleosemantics is facing as a consequence of embracing replicator ideas, it may be helpful to spell out the following central desideratum for solving the Intentional Gap:[7]

(AD) **Actuality *Desideratum***
Intentional and teleological accounts must be capable of identifying the actual, intrinsic differences between a (token) intentional/teleological system X and a (token) non-intentional/non-teleological system Y.

The idea is quite simple. Given that its naturalization is the main aim of teleosemantics, let me consider the case of intentionality as the target property of my analysis, although the argument applies to the case of teleological explanations as well. We have two different systems, X and Y. X is intentional while Y is not. Teleosemantics must be able to pinpoint, in naturalistic terms, the differences between X and Y that make the former intentional and the latter non-intentional. Which intrinsic properties of X are not present in Y? We must be able to identify not only those intrinsic properties of X that make it intentional, but also why the lack of such properties in Y makes it non-intentional. Why the movement of planets is not intentional and human behavior is? Which properties exhibit intentional systems that are not present in the movement of planets? If we are able to explain the intentional properties of X in naturalistic terms, then we bridge the Intentional Gap, we explain how intentionality is a real part of nature, and why intentional explanations are valid.

Can ET solve the Intentional Gap? It is certainly trying hard, but, I suspect, without much success. According to etiology, what are the differences between X and Y? We know that X, an intentional system, possesses a phylogenetic history guided by natural selection, while Y, a non-intentional system, does not. Saying this, however, is not enough, because we are not making explicit the intrinsic properties of X that make it intentional; unless, of course, natural selection is capable of explaining the intrinsic properties of X. What can we know about the intrinsic properties of a system just by saying that it is the result of natural selection? Certainly, not much, insofar as the explanation of this system in terms of natural selection does not look at how the inputs act on development, nor at the intrinsic properties of the outputs that make it being selected. From the evolutionary perspective of the MS, the black-box of development is also a black-box for those intrinsic properties of individual systems that the MS avoids explaining. In other

[7] As it will be pointed out later, this *desideratum* also applies to the aim of naturalizing teleology; i.e. to the closing Teleological Gap.

words, saying that an intentional system is an evolutionary adaptation does not add anything to our capacity of discerning between the two systems Y and X in terms of their intrinsic properties. Intentional systems must have different properties frm non-intentional ones, but being an adaptation does not pick out any specific property of the system that is not present in the non-intentional system.

With these remarks, I conclude that ET is facing a dilemma: either it abandons its *explanans* or it abandons its *explanandum*:

1. To abandon the *explanans* entails abandoning the strategy of rooting intentionality in evolutionary history. The AD demands an analysis of intentionality that looks into the intrinsic properties of intentional systems, and the MS view is unable to get into such intrinsic properties. The *explanans* of ET does not fulfill this demand insofar as it is based on a replicator stance.

2. To abandon the *explanadum* implies that ET is not suitable for its main explanatory aim: naturalizing intentionality. If etiology still hopes to solve the Intentional Gap, then it must show that etiological functions can account for the intrinsic properties of intentional systems. But here is where the Phylogeny Fallacy enters the scene: to accept that the evolution of the population can inform the individual/intrinsic properties of systems is tantamount to committing the Phylogeny Fallacy.

In sum, ET must abandon the *explanandum* to avoid the Phylogeny Fallacy, or it must abandon its *explanans* if it hopes to fulfill the AD. The tension I wish to highlight is between what etiology actually explains and what it claims to explain: the explanatory strategy of etiology does not match its explanatory aim.

Somehow my reflections are connected to the story of Swampman. Swampman is a replica of a human being that materializes through the action of a sudden random process (e.g. a random collision of atoms). The human being and Swampman are identical systems, both behaviorally and physiologically. However, Swampman has no evolutionary history; therefore, it has no intentionality. The critical point is not only that both systems are physiologically identical, but that ET has no explanatory tools to account for their intentional differences. This is so because etiology cannot get under the skin of the system and analyze its intrinsic properties. If we encounter a certain system and we do not know whether it is intentional or not, etiology can only tell us if it has evolved. However, intentionality is a property adjudicated to particular systems (tokens); it is a property of cognitive systems, not of populations (types). It is 'right there', causing and being caused by regularities. The explanatory inadequacy of ET, due to its populational and replicator grounds, is that it is not a theory that looks 'right there'.

Let's now turn to Shea's take on ET (Shea, 2007b, 2011, 2012a, 2013). There are two particularities in his view. First, he develops a teleosemantic theory that

not only applies to the cognitive level but also embraces the whole biological realm. Second, he accepts many non-neo-Darwinian ideas. Particularly, he endorses the Parity Thesis (other developmental resources beyond genes are equally important to explain development) and extended inheritance systems.

He called his view Infotel, since it combines informational semantics and teleofunctions. Regarding the informational side, he models development with the tools of informational theory, analyzing causal relations in terms of sender-receiver systems that produce an output. Inputs are inherited units that need not be genetic. The consumer is the developmental system itself, and the output is the phenotypic product. As for the teleological side, Shea is an etiologist (even though, as noted in Section 3.3.1, in Shea (2018), he accepts other sources for functions). However, he circumscribes etiological functions only to those systems that transmit inherited information to development and that have been selected for transmitting information to development. With this move, he opens the way to understand inherited information teleologically. Therefore, since informational inputs to development have a teleofunctional dimension (due to etiology), they are *inherited representations* (i.e. inherited information with semantic and normative properties). So the link between inputs and outputs can be understood in a teleosemantic way. Inputs represent phenotypic outcomes by being selected to transmit the information that the developmental system must *read* in order to achieve such outcomes.

There is a straightforward way to see why this is still Replicator Biology —a similar criticism can be found in Griffiths and Stotz (2013, chapter 6) and Griffiths (2013). As explained in Section 4.2.2, Replicator Biology may be assumed even accepting extended inheritance systems. Or, in other words, accepting extended inheritance systems does not entail getting into how developmental processes construct phenotypic outcomes in each generation. In a nutshell, Shea is appealing to an extended replicator view. Inherited representations are defined by the links between trait fitness and inherited units during selection processes. According to him, such a link is teleosemantic, as his Infotel theory defines it. But it says nothing about how the link between inputs and outcomes —i.e. the mechanisms of development— is traced. As noted, this is the core of replicator biology: doing evolution by looking into intra-generational inputs and phenotypic outcomes, but avoiding the analysis of the internal processing (development). Shea's view is not about developmental explanations, or about how traits function, but about how developmental mechanisms and trait functions have played a role in evolution that deserves to be treated in etiological terms. But, as asserted by Paul Griffiths: "an evolutionary explanation of a development mechanism is not the same thing as a mechanistic explanation of development" (Griffiths, 2013, 29).

So, if we are right in declaring that Shea's theory of inherited representations is a replicator account, then he is in the same complex situation as those

replicator accounts exclusively based on genetic inheritance. His *explanandum* is to understand the map from inherited information into phenotypic outcomes in teleosemantic terms. But to do so from an evolutionary perspective, he must argue that inherited representations can explain the intrinsic causal mechanisms that make developmental systems intentional. However, if this is so, evolutionary processes would be explaining individual-level phenomena, and this is an instance of the Phylogeny Fallacy. Shea's view says nothing about the intrinsic properties of development that make it intentional: "inherited information can only be used to answer 'ultimate' or evolutionary questions about inherited phenotypes, and not 'proximate' questions, such as how genes influence those phenotypes" Griffiths and Stotz (2013, 170). So Shea cannot comply with the Actuality *Desideratum* for teleosemantics unless he endorses the Phylogeny Fallacy. As Shea attempts to understand development in teleosemantic terms, the Phylogeny Fallacy also incurs in his (extended) replicator account. The moral of my analysis is that no form of the replicator stance can abide by the Actuality *Desideratum*. Genetic or not, replicator biology is silent about the intrinsic properties of teleological/intentional systems.

I would like to close this section emphasizing a marked tension in the ideas exposed so far. Here, I presented the Actuality *Desideratum* (AD). In Section 3.2.2 I introduced another desideratum for teleosemantics, namely the Historical *Desideratum*. It is not difficult to come to appreciate the tension arising between them. On the one hand, teleosemantics must be capable of explaining the current difference between intentional and non-intentional systems. On the other hand, a historical dimension is necessary, insofar as it is the history of a system (whether an individual or a population) what determines its teleofunctions. In Part III, following Waddington's 1957 taxonomy, I will argue that the developmental time-scale could be a means for alleviating this tension, because, even if it relies on proximate causes, it also possesses a historical dimension. Developmental pathways may be suitable enough for accounting for both *desiderata*.

5.3 Causal pitfalls

In this section, I present the third challenge to ET by joining together the idea of populational causation presented in Section 2.2.3, the connection between populational causation and ET introduced in Section 3.3.1, and the criticisms to populational causation presented in Section 4.3. The rejection of populational causation has three important consequences for ET: the causal epiphenomenalism of etiological functions (Section 5.3.1), the inadequacy of etiological norms (Section 5.3.2), and the need for a reconsideration of Dichotomic Thinking (Section 5.3.3). I gather these three issues under the label of *causal pitfalls*, because they they all become

manifest once etiology looses it causal roots.[8]

5.3.1 Proper functions without forces?

In Section 4.3, I argued that the Causalist School is challenged by statisticalism. Indeed, to the extent that etiologial functions are rooted in populational causes, an immediate implication for ET is that it fails on the side of its biological underpinnings. Etiological functions do not have the causal backup that is expected from a biological theory of teleofunctions.

As I already pointed out (cf. Chapter 3), etiological functions are based on selection-for effects, according to Sober's terminology: "On an etiological theory, functions are what entities were selected *for*. Mere selection *of* a trait is not enough to confer a function on it" (Neander, 2017b, 132, emphasis in the original). Selection-for effects refer to the causal role that trait types have played during natural evolution that makes such traits being selected. Etiological functions are about a causal notion concerning populational entities (trait types). Etiological functions, thus, have their causal underpinnings in populational causation. The connection between an evolutionary selected function and populational causation lies in the fact that

> election does more than merely distribute genotypes and phenotypes [...] *by* distributing existing genotypes and phenotypes it plays a crucial causal role in determining which new genotypes and phenotypes arise. (Neander, 1995a, 585; emphasis in the original) (s)

We can easily recognize a preliminary problem. Etiological functions cannot be selected-effect functions —as Sober defines them— because, under a statisticalist reading, there are no such things as selection-effect functions *qua* causal functions. The causal work is done by trait tokens, not by trait types.

Besides this point, a deeper issue has to do with the fact that ET appears to fail in its main aim: solving Brentano's Problem. Let's recapitulate. The Teleosemantics' core is to solve Brentano's Problem by looking into a biological solution to Kant's Puzzle. In a nutshell then, the problem is that insofar as the MS solution to Kant's Puzzle is misguided, and ET rests on such a solution, ET is mistaken too. Remember that the issue is about causation. The causal grounds of teleological functions act as the naturalist land to anchor intentional causation. However, the etiological teleosemantic project would not grow if such land is polluted. The causal grounds of intentional explanations are absent if they are rooted in populational causation. In other words, a naturalistic project would not be achieved if biology provides the wrong answers.

[8] This section is an elaboration of material presented in Rama (in press).

The lack of causal grounds for etiological functions is a deep theoretical problem in ET. The statistical underpinnings of etiological functions appear to militate against the purposes of etiological teleosemanticists. I believe that the correct interpretation of evolutionary functions is in statistical terms. Etiological functions are statistical functions. They have to do with the increase in fitness that a trait *type* has had during evolution thanks to having specific functions. Such change in fitness is trait fitness: it is an average calculated within a population on the basis of the Darwinian fitness of each and every individual in that population.

One could try to rescue etiological functions and suggest that maybe etiological functions, *qua* statistical functions, are still valid for a teleosemantic project. This project does not seem to work for three reasons.

First, an important reason to be developed in the next subsections is that etiological teleosemanticists themselves reject a statistical theory of functions as suitable for teleosemantics. A second, important point was already pointed out: a statistical theory of functions cannot account for the causal basis of teleofunctions. The burden of explaining the causal asymmetry on teleological explanations must carried by causal theory of functions. If populational forces exist, as the Causalist School assumes, then such causal grounds would tackle the causal asymmetry of teleological explanations. But this is taken to be impossible given a statistical reading of natural selection. So, a statisticalist theory of functions seems not to solve the riddles we want to unravel. Finally, we have a third reason why a statistical theory of functions seems inadequate. As Garson and Papineau explain:

> functions are explanatory. One peculiar feature of functions is that, when biologists attribute a function to a trait, they are often trying to give *a causal explanation for why that trait exists*. One virtue of the selected effects theory is that it makes sense of this explanatory aspect of functions. (Garson & Papineau, 2019, 36; emphasis addeed)

That is an ontological point. Teleology is connected with the existence of a trait in nature. As noted, the aptness of a trait is, in part, a consequence of having certain teleological functions. Moreover, the aptness of a trait explains why such a trait is present, maintained, and spread in nature. Therefore, the causal role a function has is crucial to explain why the trait exists, and the other way around, the absence of such a function could be a reason for the extinction of certain traits. Garson and Papineau argue, as most etiologists do, that etiological functions, if understood in causal terms, are able to fulfill the requirement that any theory of teleofunctions must account for. If we now move to representational systems, we notice that their teleofunctions must be connected with the causal role that they play or have played and that explains their presence in nature. In ET, such causal role is performed by natural selection as a populational process

dealing with representational types. The problem, therefore, is that a statistical theory of functions appears not be adequate to fulfill this ontological requirement. Even though populational/statistical explanations are related to the explanation of why certain traits are present in nature, the causal explanations of such issues cannot reside at the populational level insofar as the populational level has an epistemological status, not an ontological one.

My proposal here is an extension of Bickhard's critique of ET. Bickhard (2003) emphasizes the causal epiphenomenalism of etiological functions at the individual level. Given that etiological functions are defined as populational functions, they are separated from individual processes (something that we already saw in the discussion of Replicator Biology and teleosemantics). The definition of function is divorced from what a particular trait does. In this sense, Bickhard argues, etiological functions are causally epiphenomenal concerning trait tokens. Here I propose an extension of this idea. Etiological functions are also causally epiphenomenal at the populational level. There are no such causal roots in populational biology to anchor etiological functions. Etiological functions, therefore, as I argued, should be understood in statistical terms. They are causally epiphenomenal at any level of analysis. They are not causal but statistical.

5.3.2 A statistical theory for non-statistical norms?

The second challenge against the causal roots of etiology concerns normativity. It is a direct consequence of the first challenge regarding etiological functions. Insofar as the natural norms of traits are associated with their proper functions, the problems in the latter extend to the problems in the former. In other words, without a proper theory of natural norms, we cannot solve the Causal Mismatch, and to have such a theory, we first need an adequate theory of functions. The problem to be stressed in this section is the following: if we adopt a statisticalist position, etiological, teleosemantic norms are not the kind of norms that teleosemantics needs. The structure of my argument will take the following form:

1. Teleosemantic norms cannot be statistical;
2. but etiological functions, under the statistical view, are statistical functions;
3. then etiological functions provide statistical norms.
4. Therefore, etiological norms are not teleosemantic norms —i.e. they are not the norms that a teleosemantic project can appeal to.

The main point is stated in Item 1. I will provide two reasons to believe why it is true.

The first reason is quite straightforward and we just need to check the literature on ET to appreciate the inadequacy of etiological norms. A crucial prerequisite of teleosemantics, suggested by ET, is that teleosemantic norms cannot be statistical. As Neander explicitly stated: "[i]t might help to note that the normativity of biological functions is neither simply evaluative or statistical" (Neander, 1995b, 111). A genuine theory of natural norms has to do more than just identifying statistical properties of traits: "[t]he description of the normal system as the system that functions 'as designed' is thus not merely a generalization but a useful generalization in ways that surpass mere statistical generalization" (Neander, 2017a, 1161).

Probably, Ruth Millikan is the one who has expressed this requirement in the most transparent way. Indeed, she proposes a terminological distinction between *Normal* and *normal* (cf. for instance, Millikan, 1984, Chs. 1 and 2, and Millikan, 2017, Ch. 6). She "capitalize[s] *Normal*—to distinguish it from *normal* in the sense of *average*" (Millikan, 1984, 34; emphasis in the original). Or, in a wording not too different from Neander's: "[p]roper functions do not concern norms in any evaluative or prescriptive sense. They do not concern norms in a statistical sense either. On the contrary, many items usually fail to perform their proper functions" (Millikan, 2000a, 88). The distinction introduced by Millikan aims at separating what is genuine, natural normativity from mere statistical, quantitative analysis.

From these quotes, we can conclude that etiological teleosemanticists themselves believe that a statistical theory of norms is not appropriate for teleosemantics. From their point of view, etiological functions (understood in causal terms) are suitable candidates for teleosemantics insofar as they provide more than just statistical norms. This issue usually arises when discussing Cummins' theory of functions. As it is usually accepted, the lack of a teleological dimension in Cummins-functions allows Cummins to offer a statistical account of normativity (or another extrinsic parameter, such as research intentions). This is considered as an advantage of etiological functions over Cummins-functions. If we need to go beyond statistical analysis, and Cummins-functions cannot, while etiological functions can, we have good reasons to choose the latter over the former. This argument is developed in the following excerpt from Macdonald and Papineau (2006):

> By contrast [with the etiologists], all that the systems account [i.e. Cummins-functions] can offer is a statistical criterion: in *most* systems of a certain kind this kind of trait does F, so here the trait is malfunctioning in not doing F. By contrast with the etiological analysis, this statistical systems account seems to lack any normative content: it doesn't seem to show that a trait in any sense *ought* to be doing F; it just says it *isn't* doing F, and so is statistically unusual, but nothing

more. (Macdonald & Papineau, 2006, 11–12; emphasis in the original)

There is a counterargument to my argument. First, note that when etiologists reject statistical norms they are also rejecting those approaches based on typicality or high frequency. That is, the idea that something is normal just because it is usual. Etiologists deny that typicality can be the mark of proper functioning. As Neander illustrates:

> There is no incoherence in the idea that functional impairment could become typical in a population for a time, in a pandemic or due to an environmental disaster. The relevant function-dysfunction distinction does not seem to be simply the typical-atypical or expected-unexpected activity distinction. This much is fairly uncontroversial. (Neander, 2017a, 1152)

Neander's example shows that high-frequency does not match with proper functions, and hence that proper functions cannot be established in terms of quantitative analysis. This argument works at the evolutionary scale as well. Take sperm as an example. The point is that sperm need not perform its function very frequently for us to be able to say that such function is the proper one. That is, the fact that sperm only seldom succeeds in fertilizing an ovum in the course of evolutionary processes does not entail that fertilization cannot be taken to be its proper function. High frequency does not entail proper functioning just as low frequency does not imply malfunction.

Based on Neander's example, I will use the expression *typical-norms* to refer to those norms based on high frequency. To ward off my argument, then, one could reason that etiological norms, even if statistical norms, are not typical-norms. That is, if something is normal according to the etiological/statistical criterion, then it need not be typically normal too, or the other way around, that is, even a trait with a low frequency could have been selected for during evolution. Trait fitness, even if it is a statistical notion, is not related to high frequency.

This counterargument does not cancel point 2 but point 1 of my argument. As etiologists only deny typical-norms but not a statistical account of normativity, point 1 in my argument must be amended. Its formulation should be this:

1*. Teleosemantic norms cannot be typical norms.

So, unless there are further reasons for defending point 1 (as opposed to 1*), my argument does hold.

There is however a second reason to support point 1 hopefully capable of neutralizing this counterargument. I believe that there is an important reason why teleosemantic norms cannot be statistical. It concerns the role that selection-effect

functions —as populational forces— play in solving the Causal Mismatch. That teleosemantic norms must be considered in causal terms is crucial for an account of misrepresentation. As explained in Section 3.2.2, misrepresentation is understood as a mismatch between the evolved trait type and the individual trait token. Normativity is determined at the populational, evolutionary level, not at the individual, ontogenetic one. The explanatory goal of a teleosemantic project is to account for how reference and representation are causally connected and allow for the possibility of error. If such causal basis could be provided, and error is allowed, we would have a working non-eliminativist, naturalist project. This is a central advantage of teleosemantics. How does ET account for such an explanatory goal? If we accept a causalist reading of etiological functions, ET can explain the causal connections that exist between reference and representations. Such causal connections are responsible for determining semantic norms, and they concern those causal roles possessed by a representational system that made it being selected. Past reference-representation causal connections, defined at the populational level, set up semantic norms while allowing the possibility of error at the individual level. Therefore, as expected, if etiology lacks such casual grounds, it seems that the very core of the naturalization project cannot be completed.

The connection with the point presented in the previous subsection is quite clear. Teleosemantics has always been considered a suitable naturalistic project because of its attempts to anchor intentional causation in a solid scientific land. Notwithstanding, the lack of causal grounds at the populational level blows up the naturalistic aspirations of ET, and its theory of semantic norms too. We thereby have good reasons to support point 1 of my argument: to naturalize intentionality we need to provide a causal account of semantic norms. If this is so, if such a requirement is needed for naturalizing intentionality, then point 1 stands. Therefore, I conclude that, under a statisticalist reading of natural selection and etiological functions, the semantic norms provided by etiologists are not suitable for fulfilling the aim of teleosemantics: a naturalist solution to Brentano's Problem.

5.3.3 Rethinking Dichotomic Thinking

In Section 5.2, I introduced the Phylogeny Fallacy. Here I would like to touch on another point connected with this issue. The Phylogeny Fallacy does not depend on taking a specific stance in the causalists vs. statisticalists debate. Both causalists and statisticalists promote a division of explanatory labor. One of the main differences is that the Causalist School argues that answering how-questions (at the individual level) is explanatory irrelevant for evolution, while statisticalism defends that both levels, the individual and populational, are explanatory necessary for understanding evolution. Be that as it may, the Phylogeny Fallacy is about a conflation of the explanatory scope of evolutionary biology. So it could

arise either within a statisticalist or within a causalist view of natural selection. However, my point here is that a statisticalist view about the Phylogeny Fallacy encourages a reflection concerning Dichotomic Thinking which paves the way to (i) specify the kind of information conveyed by dichotomic terms, and (ii) re-defining the Phylogeny Fallacy.

I already introduced Keller's analysis of the gestation of Dichotomic Thinking; i.e. of the gestation of a "space between nature and nurture" (Keller, 2010). As she explains, the gestation of Dichotomic Thinking involved an epistemological shift: from asking how-question to asking how-much questions. How-questions are about how a trait is produced during ontogeny. How-much questions are about how much a trait type varies in a population under different developmental conditions. What Keller shows is that this epistemological shift involves changing the object of analysis: how-questions deal with *individual traits*, while how-much questions deal with *trait differences within a population*. How-question are about the causal mechanisms of development that produce an outcome. How-much questions aim to answer how much a trait depends on nature or nurture. To do so, how-much questions look at the difference between traits in a population to track the degree of variability that arises in different developmental scenarios.

With this analysis, Keller argues that the problem of Dichotomic Thinking is that it tries to answer how-question by providing how-much answers. In other words, we cannot know how traits develop just by analyzing the difference between traits in a population. Certainly, we can answer how much variability there is in a population as a consequence of its members having different developmental resources. However, these answers cannot be transferred to the individual level for two reasons. Firstly, while we can say how much variability exists in a population as a consequence of having a developmental cause X, it does not make sense to ask how much X contributes to developmental processes. A trait in a population having the developmental resource X exhibiting a low degree of variability does not entail that X contributes 'a lot' to developmental processes. Secondly, and more importantly, the analysis of variance in populations does not say anything about *how* a causal factor participates in development. Dichotomic Thinking appeals to an *explanans* incapable of tackling the *explanandum*.

My point here is that Keller's argument can be linked to statisticalism. The statisticalists's *dictum* that populational explanations are not causal makes clear why populational analysis does not inform about the causes of development. Trait differences are the result of populational analyses. It is about trait types. From this populational perspective, a trait may be labeled as innate if it exhibits a low degree of variability in a population, if it is developed by individuals in a population with different environmental circumstances, or if there is a high degree of heritability —low variability across generations. In all these cases, innateness refers

to properties of trait types as the result of asking how-much questions about trait differences. Statisticalism allows us to appreciate why answering these how-much questions cannot be translated to individual causation: simply, because how-much answers are not causal but statistical. So as dichotomic terms are based on populational analysis about trait differences, the statisticalist view also allows for a re-definition of the Phylogeny Fallacy. The fallacy is about the use of evolutionary explanations to account for the idiosyncrasy of individual-level phenomena. It is an epistemological fallacy affecting the epistemological tools used to account for certain biological phenomena. The interpretation of the fallacy is usually presented from a causalist view of natural selection. Accordingly, the conflation of explanations entails that evolutionary causes are part of the explanation of how-questions. Proximate causes are replaced by ultimate causes. Under a statisticalist view, the situation is not the same, however. Following statisticalism, the conflation of explanations gives rise to a different picture: it is not a conflation of different kinds of causes, but of different kinds of explanations. The Phylogeny Fallacy is to appeal to populational/statistical explanations to deal with individual/causal phenomena.

If dichotomic terms are understood in statistical terms, are they legitimate? I think that they are not. As it was explained in Section 5.2.1, there are many problems with Dichotomic Thinking. Understanding dichotomic terms as conveying information about trait types and their differences within a population allows us to appreciate why dichotomic terms are silent about the causal paths in development. A similar question arises if we look at etiology. If etiological functions are taken as statistical functions, are they conceptually valid and explanatory useful? I think that the answer to this question is affirmative. However, we need to be cautious. Explanatory adequacy is relative to explanatory aims. A classical explanatory role of etiological functions drastically changes from a statisticalist viewpoint. Particularly, in this section, we have seen how the statisticalist viewpoint makes ET a misguided project for naturalizing intentionality.

5.4 Summary

In this chapter I presented three interrelated challenges to ET. The first one denies the Explanatory Externalism of etiology: intrinsic forces also account for the existence of traits. The second one is a consequence of the replicator-based underpinnings of ET that drive etiology to the Phylogeny Fallacy. Finally, the third one impugns the idea of anchoring intentional causation at the populational level by arguing that teleosemantic needs solid causal grounds (for naturalizing both functions and norms) and that, under a statisticalist reading, populations are not the land of evolutionary causes.

There are connections between these challenges, but also an argumental transition from the first one to the last. This transition is illustrated in Figure 5.1. In the first one, I contend that internal causes are also central in the explanation of aptness and biological functions. In this sense, I reject the idea that the explanation of functions rests entirely on a one-force model based on populational causes. In the second one, I claimed that the Phylogeny Fallacy is a consequence of positing both internal/individual and external/population explanations of the adaptive complexity of developing organisms, i.e. that the Phylogeny Fallacy takes place by adopting a two-force model. Finally, I focused on the statisticalist *dictum* that the causes of adaptive complexity lie at the individual level, so no function can be explained by populational causes insofar as there is no such thing as populational causes.

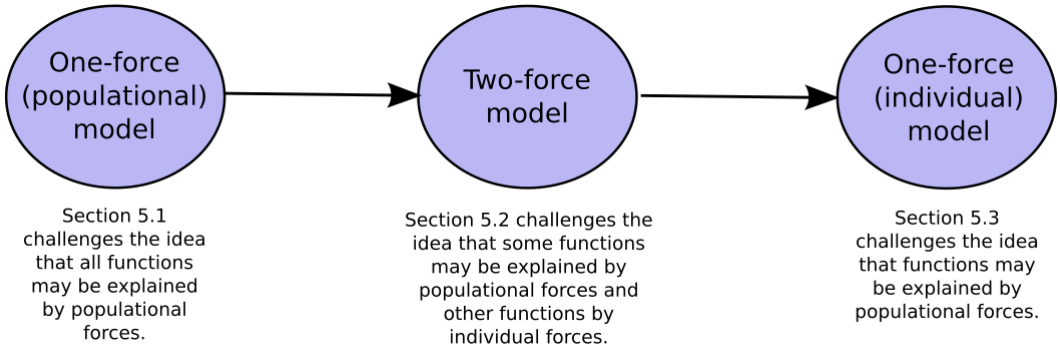

Figure 5.1: Different challenges to etiology.

As Figure 5.1 shows, I started by abandoning the one-force model —i.e. the idea that natural selection is the only adaptive force in evolution— of the MS based on external/populational causes. I continued by highlighting different problems of the two-force model and the idea that individual/internal and populational/external causes can be joined to explain the properties of individuals developing and organisms functioning. I concluded by defending a one-force model, but in this case, entirely rooted at the individual level —i.e. individual development is the only adaptive force in evolution. So the argumental connection is a transition from a one-force/populational model to a one-force/individual model, through a two-force model.

The challenges posed to ET are about etiology, not teleosemantics itself. The possibility of a non-etiological teleosemantics still remains open. Neander's book is a clear manifestation of this. Even though she was one of the main defenders of etiology, she tried to set up the core of teleosemantics and to show why it is a

prominent attempt to naturalize content beyond any commitment to a particular theory of teleofunctions. The criticisms presented here teach us important lessons for developing an alternative, non-etiological teleosemantics. I will pursue this goal in the third part with what I shall call *Agential Teleosemantics*.

Part III

New answers

Chapter 6

Teleological agency at the physiological scale

> La vie est l'ensemble de fonctions qui résistent à la mort.
>
> Xavier Bichat, 1805

6.1 The physiological level

Once the MS view on teleology has been challenged, it is time to provide an alternative view of natural teleology. In Section 1.1.2 I presented Waddington's tripartite division of temporal scales in biology: the evolutionary time-scale, the physiological time scale, and the developmental time scale. I argued at length, in Chapter 4, on the basis of contemporary debates in the philosophy and theory of biology, that the evolutionary time-scale is not the suitable ground for anchoring natural teleology. I pointed out different problems in the MS' solution which are connected with its foundational pillars. The moral should be clear by now. To the extent that teleological explanations are required to understand the aptness of living beings, and natural selection (the core of the MS' solution) is not the place to seek teleological explanations, then natural selection cannot account for the aptness of living beings. To be more specific: this does not mean that the theory of natural selection is not explanatorily important for understanding aptness, it rather means that natural selection is not capable of providing the full picture about aptness. There is a central missing ingredient in the MS view: causes. As defended by the statisticalist view of natural selection, the causes of adaptive evolution do not lie at the evolutionary, populational level. The causes of aptness are central in any naturalist theory of teleology. As noted, such naturalistic theory

must explain how teleological explanations can be understood in causal terms. As natural selection does not provide the causes of adaptive evolution, it seems to be incapable of dealing with Kant's Puzzle, i.e. explaining the causal ground of teleology. If we still want to try to look for a naturalistic theory of teleology in biology, we must look elsewhere.

Following statisticalism, it is quite clear where teleological explanations must be rooted. As statisticalism puts the causes of adaptive evolution at the individual level, teleological explanations somehow must be grounded in the individual level. Crucially, at the individual level is where are located the other two of Waddingtons's three time-scales: the physiological one and the developmental one. So if a spark of a natural teleology is to be found, it must be sought in one of these time scales, or both. This chapter is a defense of teleological explanations at the physiological time scale. The key concept is agentivity. I will argue that the agential capacities of autonomous organisms are the key to understanding teleology in naturalistic terms. *Agential teleology* is, therefore, a naturalistic attempt to legitimize teleology in biology by looking into the agentive capacities of individuals.

As Kant's Puzzle is about the temporal dimension of teleological explanations, we must evaluate at which level (if any or all) teleology can be naturalized. This chapter is about the physiological level and it is structured as follows. I will start this section by delineating the different positions with respect to Kant's Puzzle that have been defended. This will help me to locate Kant in relation with each position, as well as to specify his influence in teleological physiology —i.e. the analysis of teleology at the physiological time scale. Next (Section 6.2), I will single out the historical precursors of the contemporary accounts on agential teleology at the physiological scale. Once the historical path towards the present has been traced, I will present a detailed contemporary view on agentivity and teleology based on autonomous systems theory (Section 6.3). I will conclude (Section 6.4) by connecting the view on autonomous agency with a natural account of teleology, functions, and normativity. Most of the ideas presented here have been developed by different scholars in the last decades. While I believe that autonomous systems theory does succeed in its attempt to naturalize teleology at the physiological scale, I will point out three problems of this account: first, autonomous system theory does not explain how autonomy is constructed, secondly, it has no historical dimension, and finally, other teleological processes in nature remain unexplained within this framework. Therefore, a view of teleological development is central to act as a bridge between the physiological and the evolutionary scales. The developmental scale will be the topic of the next chapter.

Philosophically, naturalistic accounts of teleological physiology are usually tied to organicism. Biologically, the main property is that of self-organization. Self-organization, among other related properties, allows us to appreciate the agentive

capacities of living beings. Organicism has always pursued a view of living beings as goal-directed systems, i.e. as agents pursuing goals. Kant is, for one, the main precursor of organicism. However, it is important to refresh the reader's memory about his position concerning natural teleology in order to discuss some current proposals later. His position can be summarized with following three points:

(i) Kant was not a realist about teleology.

(ii) However, he believed that teleological explanations were inevitable to understanding living beings.

(iii) But, he was also convinced that teleological explanations, even if inevitable, cannot find their place in natural science.

Starting with the last point, the main problem with teleological explanations is backwards causation. This brakes the Causal Asymmetry Principle (cf. Section 1.1) that is central in post-Newtonian science. Teleological explanations cannot find a place in the science of mechanisms. However, concerning point (ii), Kant saw teleological explanations as inevitable. But this was not a demand of nature but a "demand of reason" (Desmond & Huneman, 2020). That is, for Kant, human understanding inevitability entails seeing organismal activity as goal-directed. The need for teleological explanations is imposed by our faculty of knowledge. It is in this sense that he was not a realist. Teleological explanations are not *constitutive* of nature, but just *regulative* for science: "Kant sees the capacity of teleological judgment to serve as an important heuristic device for our understanding of the natural world" (Lotfi, 2010, 125).[1]

In Chapters 1 and 3, I identified three attitudes towards Kant' Puzzle: eliminativism, anti-naturalism, and naturalism. They are the classical stances that arise when answering the following questions:

Ontological question: are purposes a real part of nature?

Epistemological question: can teleological explanations be reduced to non-teleological explanations?

Scientific question: can teleological explanations be scientifically tractable?

Eliminativists are ontological anti-realists and epistemological reductionists, and thus accept the scientific tractability of teleological language insofar as teleological language is just a shorthand for mechanical explanations. Even though

[1] But see Quarfood (2004, 2006) for a critical analysis of the regulative/constitutive distinction in Kant's philosophy.

teleological explanations could be epistemologically useful for heuristic reasons, they can be reduced to non-teleological explanations. Teleology is just 'as-if' talk: we treat organisms as if they were purposive systems just because this could help us guide our scientific research —a classical defense of eliminativism can be found in Nagel (1979).

Anti-naturalists are ontological realists and epistemological anti-reductionists, but they do not believe that teleological explanations can be scientifically tractable. Anti-naturalism is routinely associated with some form of vitalism. The core idea is that natural purposes are a real part of nature and, thereby, teleological explanations cannot be reduced to non-teleological explanations. Classically, vitalists saw natural purposes as a distinct kind of nature in the universe, a different force that is present only in living beings as, for example, the so-called living forces or Driesch's *élan vital*. However, although most vitalists were scientists (paradigmatically, Driesch (1908)), they believed that vital forces cannot be under the scope of science. However, there may be non-naturalist positions beyond vitalism. For instance, the theological proposal of Paley postulates the existence of real purposive entities such as souls but he does not attempt to treat souls from a scientific viewpoint. So both Paley's theological account and vitalism stand in the same position concerning the three aforementioned questions (ontological, epistemological, and scientific). However, the difference lies in their motivation. While Paley's anti-naturalist stance was motivated by the existence of an intelligent designer (God), vitalism was motivated by scientific, principally embryological, studies and the belief that embryology is not capable of explaining the goal-directedness of development.

Finally, naturalists are ontological realists and epistemological anti-reductionists who believe that teleological explanations must be treated as genuine and valid scientific explanations. This position differs from anti-naturalism just on the idea that there could in fact be a science of purpose. So it is expected that with the advance in science, naturalization will ensue. This is indeed what we can see if we look into the history of biology. Naturalism concerning teleology is mostly a contemporary stance, while anti-naturalism is typically a much older position. Some accounts of teleology within the MS framework are a direct consequence of the progress of evolutionary biology since Darwin. In this section, we will see how different improvements in science in the previous century enabled a different, non-Darwinian naturalist project. Table 6.1 on the next page summarizes the three position in relation to the three central questions.

	Ontological question	Epistemological question	Scientific question
Eliminativism	No (K)	Yes	Yes
Naturalism	Yes	No (K)	Yes
Non-Naturalism	Yes	No	No (K)

Table 6.1: Typical answers to the central questions around Kant's Puzzle; the letter *K* singles out those answers supported by Kant

We can easily identify the three positions (eliminativism, naturalism, and anti-naturalism) at the evolutionary time scale. There are, within the MS stance, different eliminativist accounts of natural teleology. Clear examples of an eliminativist stance are Rosenberg (2011)'s emphasis on *conspiracy theorists* behind teleological talk in biology, or Godfrey-Smith (2009)'s view of teleological talk as a consequence of humans tendency to project their own psychological capacities into the living world; i.e. as a psychological habit. As remarked, many evolutionary biologists defend that what Darwin's theory showed is that teleological and functional language could be reduced to the mechanism of natural selection. At the evolutionary level, there are also non-naturalist accounts, although in this case not within the scope of the MS but of other accounts, such as that of Intelligent Design. Finally, the MS also allowed for the naturalization of a teleology based on evolutionary processes. Natural Teleology at the evolutionary scale was largely introduced in Chapter 2 and Chapter 4 with the ideas of evolutionary design, proper functions, selection-for, and genetic programs, among others, that allowed biologists to build a functional and teleological talk from an evolutionary perspective.

At the physiological time scale, we also find the three positions. Interestingly, Kant cannot be included in any of them. This can be appreciated by taking a look at Table 6.1. He was an "ontological reductionist [anti-realist concerning natural purposes] and an epistemological anti-reductionist [concerning teleological explanations]" (A. Weber & Varela, 2002, 108). So, Kant was not an anti-naturalist because he was not a realist, nor was he an eliminativist to the extent that he was not an epistemological reductionist. Moreover, he was not a naturalist either, because he did not believe that teleology could be scientifically tractable —there will never be a "Newton of the Blade of Grass".[2] Interestingly, as we can also appreciate in Table 6.1, Kant has nonetheless motivated the three different positions up

[2]Zammito (2006) correctly points out that naturalism is a contemporary perspective in science, so it doesn't make much sense to ask whether Kant was or was not a naturalist concerning teleology. Of course, this does not prevent us from wondering whether neo-Kantians are naturalists or not.

to the contemporary accounts (Gambarotto & Nahas, 2022; A. Weber & Varela, 2002). This is so because, insofar as he did not share *all the answers* with any of the three positions, he did support *some answers* within each position. Siding with eliminativism, he was not a realist; like naturalists, he was an epistemological reductionist; and, like anti-naturalists, he did not believe that teleology could be treated scientifically.

Kant, therefore, provided important insights into the different positions concerning teleology at the physiological time scale. Along with earlier vitalists like Driesch or Bergson (Bergson, 1907; Driesch, 1908) and the German post-Kantian romantic *Naturphilosophen* (Richards, 2002), today there still survive certain non-naturalist positions, such as the one pursued by Desmond and Huneman (2020). Like Kant, they argue that teleological explanations are a 'demand of reason'. Desmond and Huneman are thus epistemological anti-reductionists, but heuristic concerning the scientific value of teleological talk —i.e. teleological talk helps in scientific research but it does not provide genuine scientific explanations (Gambarotto & Nahas, 2022). As for eliminativism, Kant's endorsement of the Newtonian scientific-style was also taken as an insight into the post-Kantian mechanistic view of the world by providing a "reductionist-leaning reading that has been most influential in the Anglo-Saxon world" (A. Weber & Varela, 2002, 103). Nowadays, for instance, Cummins-functions —defined at the physiological time scale— presuppose an eliminativist position. Like many physiologists, Cummins posited that the functioning of a system can be analyzed by observing the interacting parts without introducing any teleological language (Cummins, 1975). Finally, Kant's insight also motivated naturalistic accounts of natural teleology, and organicism belongs in this team. This is the main topic of this chapter, so I will not enter into the details yet. The organicist tradition is usually presented as the 'third way': it is not eliminativist concerning natural teleology and agency, but it is free from the metaphysical commitments of vitalism (such as non-material vital forces); organicism thus combines "the insistence of vitalism on the real complexity of life with the heuristic virtues of the mechanistic practical attack" (Needham, 1936, 9). As pointed out by Daniel Nicholson and Richard Gawne, organicism is an attempt to keep the positive insights of vitalism and mechanicism without their negative commitments:

> The vitalists had been right to defend the autonomy of biological theory, but they had been wrong to ground it in the supposition that the characteristic features of organisms are derived from the activities of unknowable directive agencies. On the other hand, the mechanicists had been right to insist that organisms are subject to the same laws that govern the operation of non-living systems like machines, but they had been wrong to conclude that such laws suffice to account for

> the character and behaviour of organisms. (Nicholson & Gawne, 2015, 358)

With these introductory comments, we can move on and start tracing the route from Kant's ideas to the contemporary, non-transcendental, and naturalistic accounts on agential teleology at the physiological level.

6.2 The historical journey of organicism

6.2.1 Kant's legacy: self-properties and causal loops

Although Kant was not a naturalist concerning teleology, he advanced two core and interrelated ideas for a contemporary naturalistic view on agential teleology: the *self-properties* of living beings (McLaughlin, 1990), and the causal loops between efficient and final causes (i.e. between means and ends). Self-properties evidence the place of natural purposes in nature, while causal loops provide the key to understanding purpose without involving backward causation.

The first element of Kant's legacy for a contemporary naturalist account is the importance of self-properties. There are many biological phenomena central to organismal life that are brought about by the organism itself. First, organisms are self-organized. Each part of the organism is related to the other, assembling a nested hierarchy of parts interacting one with the other in different physiological and behavioral processes central for the functionality of the organism. Unlike the molecular view inspired by Schrödinger, the order and organization of the living being, according to Kant, cannot be explained only by studying sub-organismal parts; also, organisms are self-maintained entities. By transforming the resources of the environment, organisms are capable of producing the matter and energy needed to sustain themselves. From metabolism to complex physiological processes and behavior, the activities that organisms perform secure the preservation of the system.

Many other self-properties are central and some of them will be discussed later, such as self-reproduction, self-regulation, self-constraint, self-determination, and self-reference, among others. However, for Kant these phenomena could not be explained from a mechanistic, bottom-up approach. This and Kant's commitment with the Newtonian paradigm lead him to believe that teleology cannot be part of science. The core problem then is that the self-properties of organisms cannot be explained simply by analyzing how sub-organismal parts interact to produce an outcome. Self-properties appear not to be tractable from a bottom-up approach: what is at stake in self-organismal properties are global, organismic phenomena. The first point then is that Kant considered these organismal self-properties to be beyond the scope of mechanistic explanations. The second point was already

noted above: he didn't believe that such properties could be scientifically explained either. As it has often been remarked, Kant's position is a direct consequence of the state of the art of the science of his time, but today, as we shall see presently, a number of approaches have been developed that seem to open the way to a scientific treatment of self-properties.

While the first element of Kant's legacy pointed to the place where natural agency might be located, the second element is the core of the answer to how natural teleology could be made scientifically tractable. In other words, the second element of Kant's legacy is central to overcoming the temporal problem of teleological explanation. It concerns $A-B-C-A$ *chains* or causal loops. The idea is quite simple. If a chain of processes starting with A leads to another process C, and C brings about another A, then C could be considered to be both the final and the efficient cause of the chain. The notions of final and efficient cause originate with Aristotle, who defined the efficient cause as the one that produces a certain event or phenomenon, while final causes refer to the end product of a chain of processes or events. In teleological explanations, means are efficient causes, and ends are final causes. A teleological explanation "purports to explain the means by the ends" (Neander, 2018, 64); i.e. why means occur because of the end-product they will produce. The temporal problem of teleology is determining how future events (ends) can explain current phenomena (means) in causal terms without positing any sort of backward causation. $A-B-C-A$ chains are a solution to this problem: in these loops of processes, ends (C) produce new means (new A) that in turn produce a new end (C). No backward causation needs to be posited and one is not forced to accept that ends are the causes of means. The recursive character of physiological processes is at the core of the view on natural purposes motivated by Kant:

> An organized natural product is one in which every part is reciprocally both end and means. (Kant, 2007, 204)
>
> As a provisional statement I would say that a thing exists as a natural end [a purposive nature, or *Naturzweck*] *if it is* (though in a double sense) *both cause and effect of itself.* (Kant, 2007, 199, emphasis in the original)

Another ingredient must be combined with $A-B-C-A$ chains for solving the naturalization of teleology: $A-B-C-A$ chains must produce an adaptive system or be directed to an adaptive goal. That is, not any causal loop deserves to be treated in teleological terms. For instance, the hydrological cycle exhibits causal loops, but there no teleology is present because no adaptive result is obtained in the cycle. Only $A-B-C-A$ in which aptness is the central *explanadum* are relevant for our investigation. The only $A-B-C-A$ chains of interest

are those showing an adaptive bias. We will see that $A-B-C-A$ chains at the physiological scale are adaptively biased, and also are adaptively biased the $A-B-C-A$ chains in the evolutionary process by natural selection (in the MS framework), unlike the hydrological cycle where no adaptive regulation is present.

In Section 6.3 I shall explain how self-properties (where natural purposes are located) are nowadays naturalized by treating scientifically different $A-B-C-A$ chains in physiological processes, but before getting to this, let me show that the solution of the MS to Kant's Puzzle also involves $A-B-C-A$ chains.[3] In this context, an $A-B-C$ chain represents the lifespan of a single individual, from conception (A) to the production of an organized and apt adult phenotype (C). Now, through reproduction, an adult (C) produces a new A (a new organism). Therefore inheritance supplies the necessary link to get the loop for an $A-B-C-A$ chain. Recall, however, that an extra element is needed: these chains must explain aptness —i.e. $A-B-C-A$ chains must be adaptively biased— and this adaptive bias is selection. Multi-generational and cumulative $A-B-C-A$ chains biased by selection processes naturalize teleological explanation (in the MS framework).

6.2.2 The birth of organicism: the Theoretical Biology Club

The term *organicism* was coined in 1903 by Delage (1903).[4] Even though organicist ideas predate the 20th century, it was in the interwar period (in the 1920s and the 1930s) that organicism became a solid philosophical and theoretical perspective in biology. The main catalyst of organicist ideas was the Theoretical Biology Club[5], a group of British thinkers formed by Joseph Henry Woodger (1894–1981), Joseph Needham (1900–1995), Conrad Hal Waddington (1905–1975), and Dorothy Wrinch (1894–1976). Their main aim was to discuss, systematize, and propose perspectives about the nature of organisms. The Theoretical Biology Club was influenced by the thought of other thinkers also aligned with soem variety of organicism, such as Ludwig von Bertalanffy (1901–1972), Paul Alfred Weiss (1898–1989), John Scott Haldane (1860–1936), Edward Stuart Russell (1887–1954), and William Emerson (1856–1944). [6] Organicism may be seen as the first attempt to naturalize agential teleology beyond Darwinism. Let me discuss briefly the three main pillars of

[3] Mossio and Bich (2017) also show how, at both scales, the physiological and the evolutionary one, the core of the solution to teleological explanations lies in circular causation.

[4] According to Nicholson and Gawne (2015), but see Etxeberria and Umerez (2006) where it is pointed out that it was Georg Ernst Stahl who first used the term in the 18th century.

[5] See Peterson (2017) for a detailed historical analysis of the Theoretical Biology Club.

[6] Some relevant works within the organicist tradition published during the interwar period are Ritter (1919a, 1919b), E. S. Russell (1924), Woodger (1929), Haldane (1931), von Bertalanffy (1933), Needham (1936), and Weiss (1939).

organicism, which will be the focus of this chapter, while I shall leave for Chapter 7 the analysis of other areas of research also connected to organicism such as Waddington's work on epigenesis (Waddington, 1957) or von Uexküll's ethology (von Uexküll, 1926).

Following Nicholson and Gawne (2015), organicism may be characterized by the three following theses:

Organisms are conceptually central in biological theory. Unlike the gene-centered view of the MS, organicism defends an organism-centered biological theory. While the MS claims that all biological processes can be explained by supra-organismal (populational) processes acting on sub-organismal (genes) entities, organicism defends that organisms are the main explanatory unit in biology. Organisms are "the unit to which all biological concepts and laws must relate" (E. S. Russell, 1930, 173), because "[n]o part of the cell can exist in isolation from the whole; to imagine such is to create a conceptual fiction to which nothing corresponds in reality" (E. S. Russell, 1930, 82).

Organization is a ubiquitous and indispensable phenomenon to explain life. (Self-)organization is explanatory central to tackle the main *expalanandum* of biology: the aptness of living beings: "[p]ractically all vital processes are so organized that they are directed to the maintenance, production, or restoration of the wholeness of the organism" (von Bertalanffy, 1933, 8).

Biology is an autonomous science. Against reductionist stances, organicism states that biology deserves its own theoretical and experimental tools. We cannot understand living systems just by looking into their physicochemical components: "[t]he view that simply by means of a knowledge of the physics and chemistry of the materials and processes of the organism biology will become a branch of physics and chemistry, and so render a theory of the organism superfluous, is thus quite untenable" (von Bertalanffy, 1933, 35).

The three theses of organicism reveal why this approach is perceived as a third way between mechanicism and vitalism. As can be appreciated, its epistemology and its ontology are impregnated with an anti-mechanistic spirit. Concerning the epistemological aspect, organicism maintains that organisms cannot be explained just by looking into their parts. The adaptivity and purposiveness of living beings cannot be grasped just by appealing to mechanistic explanations: "the chief task of biology must be to discover the laws of biological systems to which the ingredient parts and processes are subordinate. *We regard this as the fundamental problem for modern biology*" (von Bertalanffy, 1933, 65; emphasis in the original). As for the

ontological position, its organism-centered perspective puts wholes over parts as ontologically primary. Parts work in a certain way due to the requirements of the whole organism. The function of parts is subordinated to the conditions imposed by the system. As Nicholson (2014, 354) recently expressed it, "[t]he whole can be said to be ontogenically prior to the parts because an organism, unlike a machine, is not assembled from well-defined, pre-existing components. Instead, the parts of an organism only acquire their identities *qua* parts as the whole progressively develops from an originally undifferentiated yet already integrated system". Moreover, the divorce with vitalism is also a central commitment of organicism, and not only for metaphysical reasons. Contrary to vitalist ideas, such as Driesch's embryology, organicism considers that the organizational properties of living beings are capable of accounting for that *explanandum* that lead vitalists to posit non-physical forces as *explanans*: "[t]here is no 'living substance' because the characteristic of life is the organization of substances" (von Bertalanffy, 1933, 48). Like Kant and the vitalists, however, organicists also are epistemological anti-reductionists when it comes to natural teleology: "you cannot conceive of a living organism [...] without taking into account what variously and rather loosely is called adaptiveness, purposiveness, goal-seeking and the like" (von Bertalanffy, 1969, 45). However, even though the organicist commitment to natural teleology traditionally tied organicism to vitalism (Gilbert & Sarkar, 2000), organicism openly advocates for a naturalistic philosophy of biology.

6.2.3 Closure and openness

In this subsection, with the aim of getting a better understanding of natural teleology at the physiological scale, I will briefly discuss a number important thinkers and disciplines that, since Kant, have contributed to the contemporary view on autonomous agency. In the course of the 20$^{\text{th}}$ century, two interrelated but distinct phenomena came to be acknowledged as central: the *operational closure* and the *thermodynamic openness* of living organization. As we will see, closure and openness are the core phenomena to define agency from an autonomous perspective. I will start by pointing out how the notion of (organizational) closure was forged by the likes of Humberto Maturana and Francisco Varela through their autopoietic theory within second order cybernetics. I will next introduce the thermodynamic conditions of living beings as open systems, principally focusing on the work by Ludwig von Bertalanffy on General Systems Theory. From the connection between closure and openness presented here, I will next proceed to introduce autonomous system theory in the next section.

A primary step towards the scientific understanding of closure was taken by Claude Bernard (1813–1878) in his *Leçons sur les phénomènes de la vie commune aux animaux et aux végétaux* (Bernard, 1879), where he introduced the idea of

milieu interieur (internal environment). Bernard's idea of *milieu interieur* emphasized the fact that certain systems —such as organisms— constituted an individualized and organized entity distinct from their environment. The preservation of an organized entity in the presence of changes in the external *milieu* served him to picture organisms as capable of compensating for the modification of the environment as a result of the pursuit of a specific goal. This goal is the core of the physiological account of teleology and it consists of the self-maintenance of the system. The stability of systems —through regulation and compensation— is central therefore to appreciating the individualization of an organic system.

A central insight to appreciate the organized capacity of the living system concerns the so-called Bénard Cells. Bénard Cells are chemical compounds that constitute a single and organized entity as a consequence of a reaction to certain environmental conditions. Specifically, when a certain thermal input is applied to a surface of non-organized molecules, they start to behave in a certain way and this behavior eventually gives rise to a macro-phenomenon: a single organized system distinct from its environment. This has helped to appreciate two properties of living systems that will be clear by the end of this section: as Bénard Cells, living systems are also organized entities distinct from their environment as the result a number of micro-activities that constitute a single and larger macro-phenomenon: the organism. The emergence of an organized macro-phenomenon from the behavior of micro-phenomena is central to understanding organizational closure. Bénard Cells, just as living beings, have a *milieu interieur*. However, in opposition to Bénard Cells, living organisms can compensate for environmental modifications. While Bénard Cells disappear if the thermal input ceases, living beings can regulate their interaction with the environment so as to maintain the system alive and organized in front of changing environmental circumstances. Unlike Bénard Cells, living systems can compensate for environmental changes in order to regulate their *milieu interieur*. This capacity is related to the openness of living beings. In this sense, we can appreciate that closure would not be enough to define living agency, but that the thermodynamic openness and interactive capacities of living systems are also central.

A further advance on the internal organization and preservation of a system, directly motivated by Bernard's work is due to Cannon (1929) and his notion of *homeostasis*. The term homeostasis refers to the capacity of living systems to regulate physiological processes in order the preserve the inner order when facing external fluctuations. Organisms, through homeostatic mechanisms, preserve their physiological organization even in changing environmental circumstances, which strengthens the connection between being organized and being individualized; i.e. a picture of organisms as different and separate entities from their environment through the capacity of maintaining the same organization in the presence of

environmental differences. A classical example of this is temperature regulation, where a well-functioning living being is capable of adjusting to different thermal conditions in the environment in order to preserve its inner temperature in optimal state.

The work of Cannon contributed to the gestation of one of the main disciplines responsible for understanding the physiology of self-organization, namely *cybernetics*. Research on the so-called first order cybernetics, principally by Wiener (1948) and Ashby (1956),[7] marks the birth of cybernetics and the first steps towards the consolidation of a scientific discipline devoted to studying, modeling and systematizing organized systems.[8] The key concept in cybernetics, related to circular causality, is that of *feedback*. Feedback involves the use of the output of a system as an input to the same system. For example, motility and other kinds of behavior are also inputs to sensory systems, because motor activity modifies the conditions and inputs of the said sensory systems. In this sense, feedback allows for the possibility of recursive, looped circular processes. The system receives an input, processes it, produces an output, and eventually this output operates as a new input to the system. This is a clear case of an $A - B - C - A$ chain.[9]

However, the most solid account of organizational closure, which inaugurates the second order cybernetics, is the well-known work by Humberto Maturana and Francisco Varela on autopoiesis (Maturana & Varela, 1980; Varela & Maturana, 1972).[10] According to their view,

[7]Interestingly, the notion of *purposive mechanism*, central in the teleonomic accounts of the MS, that treats genes as programs specifying developmental outcomes, is taken from Wiener's work on cybernetics (Keller, 2002, 110). However, the emphasis on Wiener is lost under the teleonomic account where no goal-seeking and self-organization are posited. Similarly, the notion of *teleonomy* endorsed by Mayr was also taken from the cybernetic studies of Pittendrigh (1958).

[8]Other important contributions, especially for their influence in the future development of neuroscience, are those by Rosenblueth et al. (1943) and McCulloch and Pitts (1943).

[9]Also influential in this context is the theoretical work of the biochemist Tibor Gánti and his Chemoton model (Gánti, 2003, the original work is from 1971), which defines the three core properties of self-organized systems: self-reproduction, self-maintenance (through metabolism), and the self-individualization (through the construction of a membrane). His view was influential in the study of both self-organized systems and the origin of life —cf. Moreno and Mossio (2015, ch. 7) for discussion.

[10]Another central figure of second order cybernetics was Heinz von Foerster (von Foerster, 2003); cf. Arnellos, Spyrou, and Darzentas (2010) for the connection of von Foerster's work with contemporary accounts on autonomy. Jean Piaget's work has also been central in the gestation of autonomous system theory because, as Mossio and Bich (2017, 1098; emphasis in the original) note, "Piaget elaborates the crucial theoretical concept of *organisational closure*" (cf., for example Piaget, 1971). Many contemporary proposals on autonomy are inspired by Piaget (e.g. Di Paolo, Barandiaran, Beaton, & Buhrmann, 2014), but, for space reasons, I will not get into the details of this theory here.

> an autopoietic machine is a machine organized (defined as a unity) as a network of processes of production (transformation and destruction) of components which: (i) through their interactions and transformations continuously regenerate and realize the network of processes (relations) that produced them; and (ii) constitute it (the machine) as a concrete unity in space in which they (the components) exist by specifying the topological domain of its realization as such a network (Maturana & Varela, 1980, 78–79).[11]

They define an autopoietic system in terms of a self-constructed network of interacting elements; i.e. as a closed system constituted by mutually dependent elements that they themselves manage to self-create each part. An autopoietic system is thus a self-sustaining entity capable of preserving its organization in a way that it self-defines as a distinct unit from its environment. Autopoiesis, therefore, concerns the capacity of an organism to self-organize and self-produce its own identity by constructing a complex unit of operational interconnected elements, which determines an individualized and distinct system from the environment. The main property of such a network of interacting elements generating circular processes of causal dependence was named 'operational closure' by Maturana and Varela, which is, for them, especially for Varela in his *Principles of Biological Autonomy* (Varela, 1979), the defining property of an autonomous system: "every autonomous system is operationally closed" (Varela, 1979, 58).

Operational closure notwithstanding, organisms are not *fully* closed, since they exchange matter and energy with the environment. The energy used by the organism and its material composition were once part of the environment. To be sure, as emphasized by von Bertalanffy (1950, 1969, 1972), organisms are *also* open systems. Organisms are thermodynamically open to the environment, in such a way that there is a constant flux of energy and matter running in and out of the system. First order cybernetics failed to recognize this openness of living beings, because "a feedback system is closed thermodynamically" (von Bertalanffy, 1969, 150), just as Maturana and Varela, who defined autonomy exclusively on the basis of the closure of organisms. In contraposition, von Bertalanffy developed his biological theory considering thermodynamic openness as the central feature of living systems —and the key to understanding natural teleology (Nicholson & Gawne, 2015). Systems biology is thus another prominent field of research that emerged during the 20th century and one that has contributed to the understanding of teleology.[12] This

[11] Note that, originally, Maturana and Varela's standpoint was fully mechanistic, and only later did Varela change his views (A. Weber & Varela, 2002)).

[12] Both cybernetics and systems biology can be considered to be part of the sciences of complexity, together with other fields such as dynamical systems theory, cellular automata, or artificial intelligence; cf. Ladyman and Wiesner (2020) for the connection of complexity science with

is central to understand how self-organization and self-maintenance are achieved. As von Bertalanffy asserted, "the character of an open system is the necessary condition for the continuous working capacity of the organism" (von Bertalanffy, 1950, 23): organisms are continuously taking energy and matter from the environment. Without organism-environment interactions, life is not possible. However, to transform these environmental inputs into suitable resources to maintain the organism alive, the organism must be functionally organized. In other words, it is the functional organization of the system that facilitates the transformation of external resources into available work to sustain the system. The paradigmatic case is metabolism: energy and matter enter the cell through the membrane where different inner processes are responsible for producing the needed ATP to be used to sustain the activity of the cell (growth, reproduction, repair, move, etc). To be able to obtain available work, many complex metabolic processes are involved, in such a way that what makes it possible to acquire work and not mere heat from environmental input is the presence of an organized and complex system responsible for the production of ATP. All in all, without closure, openness would not provide anything useful; without openness, closure would disintegrate.

Given their thermodynamic properties, living beings may be seen as dissipative systems, in the sense of Nicolis and Prigogine (1977). The core idea of what is a dissipative system was already implicit in the phenomenon of Bénard Cells discussed above, i.e. a collection of microscopic elements that compose a global pattern or configuration as a consequence of the exchange of matter and energy with the environment, such that, when the exchange is interrupted, the patterned order disappears. There are many other examples of non-living dissipative systems comparable to the one of Bénard Cells, such as tornadoes, candle flames, or oscillatory chemical reactions. When we turn to living dissipative systems, however, we observe a number of differences with non-living dissipative systems that completely set them apart; these differences will be relevant for our discussion. First, non-living dissipative structures arise *suddenly*, while living beings are constructed through an individual process of ontogenesis; second, non-living dissipative structures have no parts —i.e. all their components roughly play a similar causal role— while in living systems different parts play different causal roles; and last, but not least, non-living dissipative structures cannot regulate their environmental interaction while living organisms are typically able to adjust their coupling with the environment.

Summarizing so far, to understand self-organization and self-maintenance neither closure nor openness can be neglected. On the one hand, if we focused exclusively on closure, we would miss the interactive dimension and the thermodynamic openness of living beings, and we would therefore be unable to have access to how

systems biology and cybernetics.

organisms interact with their environment to stay alive and remain organized. If, on the other hand, we would only focus on openness, we would neglect the crucial aspect of how the exchange of matter and energy contributes to producing an organized living system. The dialectic between closure and openness is, therefore, the central aspect of living beings. Operational closure and organization are central to transforming the inputs from the environment into available work; while openness is central to preserving the organization. Closure allows for a productive interaction of the system with the environment, while openness makes possible the preservation of the organization.

Autonomous systems theory, in its contemporary formulation, posits organisms as autonomous agents, emphasizing the fundamental connection between operational closure and thermodynamic openness in order to naturalize the self-organization and self-maintenance capacities of living beings. Perhaps the first and most well-articulated attempt at the integration of closure and openness with the notion of autonomous system is due to Stuart Kauffman (Kauffman, 2000, 2003). As he famously declared —and his models of Work/Constraint cycles aim to capture— organisms live *at the edge of chaos*, that is, organisms sit in an intermediate position between being closed and thermodynamically frozen, and being open and utterly chaotic. Organisms, therefore, manage to live preserving an inner organization without actually becoming totally closed systems, while at the same time they stay thermodynamically open without ever losing their identity. Also relevant in this context are the works of the late Robert Rosen (for example, Rosen, 2000), who reinvigorated Aristotelian causes to treat organisms as inner causally efficient, but also materially dependent on the environment, entities. The work of both Kauffman and Rosen has played a critical role in the development of modern autonomous systems theory and their ideas permeate almost all research in this field. Notwithstanding, I will not offer a detailed presentation of their thought here, which is anyhow well-integrated into the research projects of contemporary researchers in this area, to which I turn in the following section.

6.3 Autonomous agents at the edge of chaos

In this section, I will present the most comprehensive attempt to naturalize teleology at the physiological time scale, within the framework of autonomous systems theory. Recall from our previous discussion that autonomous systems theory takes important insights from both cybernetics and systems biology, and from the aforementioned scientific and philosophical developments. It shares the Kantian spirit of putting the emphasis on self-properties and $A - B - C - A$ chains, and, to a certain extent, it also finds inspiration in ideas embraced by Theoretical Biology Club. I will start with a presentation of the contemporary view of autonomous

systems theory in order to spell out its connections with the notion of agency. In the next section I will make explicit the connections between autonomous agency and teleology, function and normativity at the physiological scale.

6.3.1 Autonomous systems theory

Contemporary autonomous systems theory is the result of the collaborative effort of different scholars in different places in the course of the last twenty years (for some key references see, Arnellos et al., 2010; Barandiaran, 2008; Barandiaran & Moreno, 2008; Bickhard, 2000a; Juarrero, 1999, 2009; Ruiz-Mirazo & Moreno, 2004; Ruiz-Mirazo, Peretó, & Moreno, 2004). A recent milestone in the field is Álvaro Moreno and Matteo Mossio's book *Biological Autonomy* (Moreno & Mossio, 2015), which will be the main source on which the material in this subsection is based.

"Autonomous systems are then, in this view, *organisationally closed* and *thermodynamically open*", write Moreno and Mossio (2015, 6; emphasis in the original) rehearsing a contention we've already met above. They associate the closure and openness of autonomous systems with two dimensions of biological autonomy: "the *constitutive* one, which largely determines the identity of the system; and the *interactive* one, which, far from being a mere side effect of the constitutive dimension, deals with the inherent functional interactions that the organisms must maintain with the environment" (Moreno & Mossio, 2015, xxviii; emphasis in the original). As noted, the very identity of the system lies at the level of organization; i.e. what defines the system as a single unit different from its environment is operational closure: the fact that it is an organized entity constituted by a network of mutually interrelated parts. Consequently, while matter and energy flow constantly through the interactive dimension, the organization remains stable; as it is well captured by Bernard's notion of *milieu interieur*, the system may remain unchanged even in the presence of constant environmental changes.

While the connections between closure and the constitutive dimension, and between openness and the interactive dimension have been observed elsewhere, Moreno and Mossio (2015) introduce a further distinction, based on the opposition between constraints and processes, that is characteristic of their view on autonomy. Moreover, as I will show later, their characterization of autonomous systems allows us to trace the boundary between autonomous systems and 'quasi-autonomous' systems —i.e. those systems that exhibit some of the properties of autonomous systems but not all of them.

Constraints refer to all the elements that constitute the organization of the system that remain operationally stable —even when their material basis changes— in the presence of environmental changes. More specifically, constraints refer to the organic structure and the different elements that make up an organism. The set of interrelated constraints is thereby defined independently of the external con-

ditions —constraints are preserved in different environmental scenarios— and are also independent of their material bases —the material makeup of the constraints of a system may change, but the relation between the constraints remains stable. At the same time, different continuous physicochemical *processes* take place within the organism acting on its constraints: an autonomous system is constantly processing matter and energy to sustain itself. These processes do not necessarily change its structure, they for the most part contribute to sustain it by both regulating its coupling with the environment and providing the fuel that makes the system work.

Within this conceptual framework, the authors are able spell out the connections existing between these different conceptual distinctions. On the one hand, the operational *closure* of living beings determines their *constitutive* dimension on the basis of a set of interrelated and mutually dependent *constraints*. On the other hand, the *openness* of living beings foregrounds the *interactive* dimension of living beings through the different *processes* that an autonomous system performs in order to sustain its organization; in the authors' own words, "biological systems can be shown to involve two distinct, although closely interdependent, regimes of causation: an *open* regime of thermodynamic *processes and reactions* [taking place in their interactive dimension], and a *closed* regime of dependence between components working as constraints [thus defining its *constitutive* dimension]" (Moreno & Mossio, 2015, 3, emphasis in the original).

Equipped with this conceptual apparatus, Moreno and Mossio eventually define autonomous systems on the basis of the auxiliary notion of *self-determination*. According to them, the idea of self-determination has to do with the fact that in biological organized systems the "effects of its activity contribute to determine its own conditions of existence" (Mossio & Bich, 2017, 1089). A system self-determines itself if it is organized and maintained by its own activity. The conceptual apparatus proposed by Moreno and Mossio (and their collaborators) allows them to specify what kind of self-determination autonomous systems exhibit: autonomous systems are self-determined as being self-constrained. That is, the constraints of autonomous systems are determined by the system itself. As they note, and we will see later, other systems realize operational closure, such as the hydrological cycle, but only living ones produce their own constraints. In the hydrological cycle, the "dynamics of the river are specified in particular by the conformation of the ground and its slope, which are not generated by the water cycle itself" (Mossio & Bich, 2017, 1106), whereas each part of the operational closure in a living system is the product of the activity of the system.

To conclude this subsection, let us try to pin down the traces of Kant's legacy in Moreno and Mossio's proposal. It is clear that their constitutive causal regime and their interactive causal regime are reminiscent of $A - B - C - A$ chains or

of recursive causal loops: constraints depend on each other to sustain the whole system, and the whole system enables in its turn the existence of each constraint. The flow of matter and energy in the interactive dimension also reminds us of $A - B - C - A$ chains, as it is also the case of metabolism or of Kaufman's W–C cycles. Moreover, such circular causal regimes make it possible the naturalization the self-properties of living beings: Moreno and Mossio (2015) put special emphasis on the role of self-determination, but I showed how this notion is tightly connected to other self-properties, such as self-organization, self-production and self-maintenance, discussed earlier in this chapter.

6.3.2 Naturalizing agency

In this section I shall discuss the main defining properties of agency. My aim is to show how autonomous system theory is capable of treating living beings as autonomous agents from a scientific viewpoint. A preliminary approximation to the idea of agency can be found in one of Stuart Kauffman's earlier books: "an autonomous agent is a physical system that can act on its own behalf in an environment" (Kauffman, 2000, 8); that is, an individual entity that does things that make sense to itself in order to stay alive and adaptive. Following Barandiaran, Di Paolo, and Rohde (2009), we can single out three properties, implicit in Kauffman's preliminary approximation, that a system must possess to be considered an agent.

Individualization: to define the agency of a system, it must be possible to trace a distinction between the system and its surroundings. As suggested by Kauffman's words, if a system is to act on its environment, then an intervening physical boundary must exist between them. Two relevant points deserve special attention. First, the individuality of the system should not be established by an external observer. The intrinsic forces of the system must constitute the individuality of the system, not an external stipulation. Second, as already noted above, the individualization condition does not preclude the system from interacting with the environment —e.g. by exchanging matter and energy. As Barandiaran et al. (2009, 370) put it: "the first condition for the appearance of agency is the presence of a system capable of defining its own identity as an individual and thus distinguishing itself from its surroundings; in doing so, it defines an environment in which it carries out its actions".

Interactional Asymmetry: An intuitive idea behind the notion of agency, also present in Kauffman's quote, is that of a system doing things in an environment. Agents act. Other systems —such as the planetary system or the hydrological cycle— don't do things, they perform no actions. In non-agentive

systems, the interaction between the system and the environment is not regulated by the system, there are no qualitative differences between external and internal forces: the interaction is symmetrical insofar as the system does not contribute any distinct kind of activity. Organisms, however, are capable of modulating their coupling with the environment, to promote certain kinds of system-environment interactions from within the system. Interactional Asymmetry concerns the capacity of agents to modulate their relation with their environment; for example, Barandiaran et al. (2009, 372) define interactional asymmetry "as the condition describing a system as capable of engaging in some modulations of the coupling and doing so at certain times".

Normativity: The activities of an agent —an individualized entity capable of modulating its coupling with the environment— seem not to be random but relative to the organism's intrinsic goals: agents do things that 'make sense' for themselves. This is the condition of normativity. Agents do things to achieve a certain purpose, to fulfill certain goals. The functioning of an agent is directed to a normative valuation. In other words, the activity of agents transforms the modulation of the environmental coupling into a regulation of the coupling, i.e. something is done so as to satisfy a given norm (Barandiaran et al., 2009, 370).

These three conditions or properties that a system must meet to be considered an agent should be seen as a minimal requirement and therefore not in exclusion of other significant properties that might be associated to agency. They are the three necessary and sufficient properties to define a minimal notion of agency. Others may also want to include, for instance, rationality or morality in the list of properties that are important in agency, but in this case it is clear that what is at stake is a more robust notion of agency perhaps circumscribed to animal agency. Be that as it may, let us see how these three properties are accounted for from the perspective of autonomous systems theory.

The first property in the definition of agency is that of individualization. Individualization was well recognized by Bernard through his notion of *milieu interieur*. Also, the autopoietic theory of Maturana and Varela presented before characterizes autopoietic systems as operational distinct systems from their environment. Autonomous systems theory was nurtured by their insights to explain how the individuality of autonomous systems emerges from the set of interrelated constraints that defines the constitutive dimension of living beings. As explained, autonomous systems are individualized at the constitutive dimension. Importantly, as the definition of agency demands, the individuality of an agentive system must be established by the system iself, not by an external force or observer. This demand is also well-accounted for by autonomous systems theory by recognizing that organ-

isms are self-constrained systems: the internal relation between its parts defines the identity of the system.

Concerning the second property —interactional asymmetry— autonomous systems are a clear case of a system displaying asymmetric relationships with the environment. As explained, autonomous systems are dissipative systems. Dissipative systems preserve their organization due to the constant exchange of matter and energy with the environment. However, what is characteristic of autonomous systems is that they modulate their interaction with the environment in a way that they can face up to different kinds of environmental scenarios. So an autonomous system has an asymmetrical interaction with the environment insofar as autonomous systems can do things to adjust themselves to changing external conditions.

Concerning the last defining property of agency, normativity, autonomous systems modulate environmental interactions to achieve their principal goal: self-maintenance. The teleological and normative dimensions of autonomous systems will be presented in detail in Section 6.3. The important point is that autonomous systems are capable of regulating their interaction with the environment in order to provide an adaptive response. Autonomous systems can regulate their coupling by different means, such as by internal compensation, plastic responses, motility, or different behaviors [13] Moreover, the normativity of autonomous systems is not extrinsically but intrinsically defined. This is well illustrated by the etymology of autonomy: auto(self) – nomous(norms). So normative parameters are not extrinsic; rather, normativity is defined by the very requirements of the system to maintain itself in viable conditions. All in all, autonomous systems exhibit the main three defining properties of agency: autonomous systems are individualized, their interaction with the environment is asymmetrical, and they modulate their coupling with the environment on the basis of intrinsic norms and towards adaptive outcomes. In conclusion, "an agent is an autonomous organization capable of adaptively regulating its coupling with the environment according to the norms established by its own viability conditions" (Barandiaran et al., 2009, 376).

6.3.3 Boundaries of autonomous agents

In Section 6.3.1 I advanced that autonomous system theory provides the tools for defining the boundaries between autonomous agents and what I presented as 'quasi-autonomous' agents —i.e. those systems that exhibit some of the properties of autonomous systems but certainly not all of them. In this section, I would like to devote some space to this issue, starting with those properties that separate liv-

[13] See Bich, Mossio, Ruiz-Mirazo, and Moreno (2016) for further developments on the notion of regulation within autonomous system theory.

ing systems from non-living dissipative systems reviewed in Section 6.2.3. Recall that non-living dissipative systems (Bénard Cells, tornadoes, candle flames) spontaneously arise due to specific environmental conditions. While they are individualized systems, self-organized, and in similar thermodynamic conditions as living systems, non-living dissipative systems do not construct their existence through a developmental process. Also, non-living dissipative systems typically have no differentiated parts and all their elements roughly perform the same causal role. This is one of the reasons for Moreno and Mossio (2015) to lay stress on the importance of constraints, insofar as non-living dissipative systems lack them. Recall moreover that non-living dissipative systems are extrinsically constrained: their existence crucially depends on the environmental conditions. Therefore, non-living dissipative lack two properties that, according to autonomous systems theory are characteristic of autonomous agents: they do not modulate their interaction with the environment (no interactional asymmetry) in order to maintain themselves (no normativity).[14]

But autonomous system theory allows us to go a bit beyond and eventually trace the borders between living systems and such artificial systems as robots, for example. Robots have a constitutive dimension (individualization), they modulate their interaction with the environment, and there is a normative valuation on their well or bad functioning. However, Robots differ from living organisms in ways that suggest that they cannot be classified as autonomous agents. The main difference concerns the fact that neither the normativity nor the individuality of robots is established by the system itself but rather by an external designer or engineer. In other words, the constitutive dimension of a machine is not self-determined but extrinsically determined. Even though machines are intrinsically constrained by their inner structure, such constraints are not determined by the system, and, unlike true to autonomous agents, the ontology of the parts of a robot does not depend on the system as whole. Another point to be taken into account is that even though (embedded and embodied) robots interact with the environment, they are not thermodynamically open. Specifically, while robots need to acquire energy from the environment, the preservation of their constitutive dimension does not depend on such energy: robots do not disintegrate when we turn them off. Moreover, while my entire body has been constructed and reconstructed through the transformation of matter coming from the environment, robots, once constructed, do not exchange matter with the environment (cf. Nicholson, 2013, 2018, for an analysis of the thermodynamic differences between organisms and machines). In-

[14]Of course, the fact that autonomous systems are a specific kind of dissipative system does not entail that physical self-organization is not a central explanatory element. As it was explained in Section 5.1.1, Newman's account of inherency stresses how different traits arise during evolution and ontogeny due to self-organization processes acting on the physical substrates of cellular and multi-cellular organisms.

deed, as Roli, Jaeger, and Kauffman (2022) argue, the interactive dimension of living beings imposes a critical limitation on artificial general intelligence.

On this account, we can appreciate why for example cells, unlike robots, can be treated as autonomous agents —usually they are treated as *minimal autonomous agents*. Cells are self-individualized entities separated from the environment through a permeable membrane. Their identity is preserved both at the constructive and at the interactive levels as structural and functional units. Moreover, cells regulate their interaction with the environment. They allow different chemical ingredients to get through the membrane to transform them into work. Also, the motile capacity of cells enables them to navigate their environment. The interaction with the environment is not symmetrical but asymmetrical: interactions through the membrane or through motility are enabled and performed by the cell itself. Finally, the modulation with the environment fulfills a certain norm. Such norm is that of self-maintenance. In this sense, the modulation of the coupling becomes a regulation of the coupling. Cells regulate their interactive and constitutive processes to self-preserve in adaptive conditions.

All in all, these are strong grounds to assume that autonomous systems theory is an appropriate framework to tackle the naturalization of agency: it captures the intuitive idea of agency outlined by Kauffman's preliminary approximation, it provides a robust scientific account, it may be applied to systems that *prima facie* would be treated as agents and, importantly, it also allows us to see why certain systems should not be treated as agents. Even though the clear definition of boundaries is always a complex task, to establish a viable distinction between agentive systems and non-agentive systems looks feasible from the standpoint of autonomous systems theory.

Turning now to another related issue, as it is often recognized (e.g. by Walsh, 2006a), with the thermodynamic properties of organisms in mind, we can better appreciate the parallelism existing between Kant's Puzzle and what we might call *Schrödinger's Puzzle* presented in his book *What is life?* (Schrödinger, 1944).[15] In a nutshell, Schrödinger's Puzzle poses a question on the status of living beings within the natural sciences. According to the second law of thermodynamics, the entropy in the universe tends to increase with time, order decreases as chaos grows. However, organisms appear to violate the second law: organisms reduce their entropy by taking energy and matter from the environment to produce a well-organized system, a system at the edge of chaos, but not in chaos while still alive. So as disorder in the universe gets bigger, order increases in the living realm. This illustrates the tension between the understating of the physical and living worlds

[15]Walsh refers to Schrödinger's analysis as a paradox. However, I prefer to avoid calling it *Schrödinger's Paradox* because another well-known phenomenon about a cat in a box already has this name. So I opt for the expression *Schrödinger's Puzzle* which reinforces the parallelism with Kant's Puzzle.

from a thermodynamic perspective, i.e. between the so-called fundamental laws of nature and the nature of organisms. It is in this tension that can be localized the parallelisms with Kant's puzzle: both in Kant's Puzzle and Schrödinger's Puzzle, a conflict arises between the view of nature provided by physics and chemistry and what is found in the living world. The tension is solved by recognizing that organisms, while reducing their own entropy, make chaos grow in the environment. In other words, if we consider the whole environment-organism system, and we do not focus exclusively on organisms, entropy tends to increase, and the second law is still valid. While this somehow removes the tensions, it does not yet offer a solution for how such complex, organized and negentropic systems come into existence in a universe governed by the second law. This was Schrödinger's explanatory target: to account for the complex organization of living beings and to be able to explain their thermodynamic, far-from-equilibrium properties.

We already know that the molecular answer since Watson and Crick accounts for order in the living as resulting from order in the non-living: a complex and ordered macro-molecule capable of creating the complexity of living beings. But as already noted before, from the organicist tradition to the autonomous system theory, another answer has been offered on the basis of the idea that an organismic level of analysis is indispensable, the so-called order-from-disorder strategy. In contrast to the order-from-order answer, Kauffman (2003), like many others, pursues a view where the thermodynamic properties of living beings are properties of the whole system, not of any of their parts. The far-from-equilibrium condition arises from the dynamics of the whole system to self-maintain its functioning and organization. So just as "one cannot reach into a glass of water and pick out a molecule and say 'This one is wet' " (Searle, 1992, quoted in Gilbert & Sarkar, 2000, 2), we cannot watch some intra-cellular element trying to observe the far-from-equilibrium thermodynamic conditions present in the whole cell. Order-from-disorder states, therefore, that disordered entities (sub-organismal parts) give rise to ordered entities (organisms) once the sub-organismal parts interact towards the maintenance of the whole organism.

I would like to conclude this section addressing a final, but highly critical point which, although not necessarily unsolvable, looks nonetheless like a tough nut to crack. Autonomous systems theory crucially relies on some notion of the individuality of organisms, but, as emphasized by many contemporary scholars, the notion of an individual organism is not so easy to define. This problem, and its relation with the foundations of organicism, has been the object of a recent analysis by Baedke (2018b). From an intuitive viewpoint, it is quite clear what an individual organism is supposed to be. However, once we get into the details of the reality of living beings, things get much more complicated. Several biological phenomena make the question of boundaries a really complex issue to solve. A classical

problem for individuality is that of symbiotic interactions. Many organic functions are dependent on the functioning of other organisms. So if the preservation of an organism is dependent on the preservation of another organism, then the characterization of autonomous systems somehow must include other autonomous systems as part of themselves.

The notion of *holobiont*, introduced by Adolf Meyer-Abich (Meyer-Abich, 1950), popularized and explored principally by Lynn Margulis (Margulis, 1990, 1993) and recently revitalized by Scott Gilbert (Chiu & Gilbert, 2015; Gilbert, Sapp, & Tauber, 2012; Gilbert & Tauber, 2016), aims to capture this symbiotic relationship between individuals. Holobionts are biological entities constituted by the interaction of multiple organisms, typically of different species. Holobionts are widespread in nature; even within a system with apparent clearly defined physical boundaries like a human, there might be other autonomous systems (e.g. microbiota) that are fundamental for the functionality of the whole system. Should these entities be considered part of the organism? Or should they be treated as separate organisms coexisting within the same physical boundaries? Should holobionts be treated as autonomous systems? What are the constitutive dimensions of holobionts? These questions require an answer from autonomous system theory. Even though this issue does not make autonomous theory and its organismic foundations invalid or useless, it is important to recognize that further adjustments need to be done toward a biological theory of autonomous systems in consonance with the reality of biological individuals.

6.4 Teleology, functions and norms at the physiological scale

In the previous section, I sketched a proposal for an account of minimal agency from an organismic point of view and based on autonomous systems theory. In this section, I will connect this with the issues of teleology, normativity, and function. As we saw, these three notions are interrelated. Teleology defines purposes, which establish a norm that trait functions attempt to fulfill to achieve such purposes. In the MS framework, we saw that evolutionary processes determine the purpose of each trait in relation to their fitness contribution. This provides a demarcation line between normal and abnormal functioning, and therefore also between proper functioning and improper functioning. These notions are central in teleosemantics. The purpose of a representational system determines how the system must function in relation to a norm. Such a norm is central to determining the content of a representation: this solves Brentano's Problem from a naturalist stand. In this subsection, I will offer an outline of how teleology, normativity, and

biological functions are understood from an autonomous systems perspective at the physiological time scale. Later on, I will appeal to these notions to promote a non-etiological teleosemantic project.

6.4.1 Intrinsic teleology

Autonomous systems theory and its organismic foundations promote a naturalist account of intrinsic teleology. Being naturalistic entails first a realist position concerning teleology. Teleology is not just an as-if phenomenon. It is a real part of nature. No transcendental view is defended —in spite of the Kantian traces in autonomous systems theory— or a relativist position is endorsed. Rather, teleology is immanent in living beings; teleology is something that an organism possesses, not something added by an external observer or by the human intellect. Moreover, teleological explanations cannot be reduced to a non-teleological, bottom-up account. The autonomy of the organism is accounted for by looking at the dynamics of the whole and by analyzing how each part contributes to the fulfillment of the goals and norms self-defined by the autonomous system. Finally, autonomous systems theory has a robust scientific background in cybernetic, complexity theory, and systems biology among other relevant scientific fields, in such a way that, contradicting Kant's prognosis, teleological explanations can fall under the scope of science.

Now, the most distinctive aspect of this naturalist account is that teleology is intrinsic to the organism. This opposes externalism. Therefore, intrinsic teleology necessarily contends that the processes responsible for attributing normative teleofunctions in the living realm are internal, not external to the organism.

Autonomous systems theory provides a view of teleology at the individual level by emphasizing the goal-directedness of physiological processes. As noted, the MS solution is externalist: Explanatory Externalism establishes that functions are determined by the external pressures during selection processes leading to adaptations. However, there could be externalist accounts not at the evolutionary time scale but the physiological one. Indeed, this is the case of the proposals by Babcock and McShea (2021) and McShea (2012), for example. According to the *field theory* developed by Daniel McShea, the environment provides the direction that a system —in this case an organism— must follow. The environment operates as a field that orchestrates the activity of the organism: "[g]uidance always comes from the outside, from some larger external entity in which the guided entity is embedded" (Babcock & McShea, 2021, 8). We can appreciate that this externalist position entails —as the externalism of the MS does— treating organisms as objects, not agents of their own activity. In contraposition to an autonomous systems perspective, here the organism does not modulate or regulate its constitutive or interactive processes, but it just passively responds to the guidance of the

external field. Externalist teleology, even if circumscribed to the individual level, still eliminates the agential role of inner, organismal processes.

In the view of teleology advocated by autonomous systems theory, relying on Kant's legacy, the main goal of a trait is to contribute to the self-properties of the system it belongs to. Particularly, at the physiological time scale, *self-maintenance is the principal purpose of the system*. This was clear in the case of minimal agency: a cell is a complex system capable of living at the edge of chaos. Each of its component parts contributes to the goal of maintaining the system in a complex equilibrium between being internally organized and interactively open to the environment. In minimal agents, the goal is the self-maintenance of the system at the edge of chaos: to stay not so chaotic as to lose their negentropic property, nor to be so stable as to become frozen. The result is that "for Kant as well as for [part of] modern biology [...] the natural purpose of life is to sustain life itself" (Lotfi, 2010, 127).

Therefore, we can see that the autonomous systems theory based on self-determination (and other self-properties) successfully accounts for the central *explanandum* in Kant's Puzzle, i.e. understanding how teleological explanations can be aligned with the Causal Asymmetry Principle by showing that no backward causation needs to be posited. This is done by noting how $A - B - C - A$ chains and the recursive character of physiological processes (self-)determine a self-organized and self-maintained autonomous system: "[t]his interrelation of means and goals [in recursive physiological processes] describes a circular situation: parts of an organism are there through the existence of the whole and the whole is responsible for the parts" (A. Weber & Varela, 2002, 106). Intrinsic teleology is thus causally anchored in the existing circular causal processes that constitute the existence of an autonomous system and that it must carry out in order to preserve its own existence. In Matteo Mossio and Leonardo Bich's words:

> Biological organisation can be legitimately conceived of as an intrinsically teleological causal regime. The core of the argument consists in establishing a connection between organisation and teleology through the concept of self-determination: biological organisation determines itself in the sense that the effects of its activity contribute to determine its own conditions of existence. (Mossio & Bich, 2017, 1089)

6.4.2 Organizational functions

Autonomous systems theory and its organismic foundations suggest a different view on functions than the one championed by the Modern Evolutionary Synthesis. This view has come to be known as the *organizational* (or, also, *systemic*) theory of functions, whose main supporters are Schlosser (1998), Kauffman (2000,

2019), Christensen and Bickhard (2002), Mossio et al. (2009), McLaughlin (2000), Saborido, Mossio, and Moreno (2011), and Moreno and Mossio (2015).

Functions are related to goals. This is the reason why teleosemantics appeals to natural teleology to specify the proper functions of representational systems. Since the principal goal of an autonomous agent is that of self-maintenance, the function of a trait is to contribute to self-maintenance; therefore, in autonomous systems a trait is "functional if it contributes to its self-maintenance" (Barandiaran & Moreno, 2008, 329). Another way of expressing this idea is by noting, as Okasha (2016) did, that goals are adjudicated to the whole system while functions are adjudicated to the parts of the system. This idea is connected to the 'unity of type' in teleological explanations: the idea that teleological explanations apply to a unified system and not to its parts. That said, if functions are related to goals, and the primary goal is that of self-maintenance, then the function of a trait must reside in its contribution towards the fulfillment of such a goal. Therefore, from an organizational viewpoint, the function of a trait cannot be individuated independently of the whole organism. A trait has its role in the system, and functions are determined by such role. Traits belong to complex physiological processes and their functioning is tied to their causal role in such processes, so "to ascribe functions we must distinguish between different causal roles in the system, a division of labor among the parts" (Moreno & Mossio, 2015, 72). In this sense, a trait must function according to its connection with the rest of the traits towards the self-maintenance of the organism (cf. Moreno & Mossio, 2015; Mossio et al., 2009, for a formal definition of organizational function aligned with this idea). Trait functions are part of the recursivity of physiological processes: a trait is needed to self-maintain the organism, and the organism secures the existence of the trait:

> 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. (Mossio et al., 2009, 828)

The impact of organicism is also remarkable in contemporary organizational accounts on functions. Trait functions cannot be designated without taking into account the whole organism. In this sense, the whole organism is the proper unit of analysis. As organicism claims, organisms, and not their parts, are the fundamental units of organization of the living. The function of a trait, therefore, is tutored by the functioning of the whole organism and its current needs: "[t]o

discover the function, we must study the whole organism in its environment. There is an unavoidable holism to biology" (Kauffman, 2003, 1097).

Organizational functions can be easily applied to minimal autonomous agents, such as unicellular organisms. Cells have structurally and functionally different parts that must interact in a certain way to self-sustain the cell. Each part must perform its proper function to maintain the system alive. In the case of metabolism, the membrane has its own proper function for metabolism to take place successfully. If the membrane lets certain toxins into the cell, this will have maladaptive consequences and it will not be functioning according to the needs of the whole organism. Similarly, if the reaction through inner chemical pathways does not generate the necessary metabolic products, the cell would not produce enough work to self-repair, act, and reproduce, it would not be functioning properly. Moreover, an organizational account of functions also allows us to see why organizational functions are only present in autonomous agents (even though boundaries are not always easy to define; cf. Mossio & Bich, 2017). Let's consider another, non-living dissipative system, such as a tornado. We already know (Section 6.3.3) that they should not be treated as autonomous agents because they cannot modulate their interaction with the environment; but neither should they be ascribed any functions, for the simple reason that a tornado has no different parts that contribute in different ways to its self-preservation. That is, there is no distribution of causal roles among its parts. Non-living dissipative systems are usually macro-phenomena that emerge due to the dynamics of micro-elements performing a similar activity.

We can appreciate how organizational functions successfully account for a central *explanandum* that a theory of functions must address, namely, the reason why a trait is present in nature. Cummins characterized this explanatory aim thus: "The point of functional characterization in science is to explain the presence of the item (organism, mechanism, process, or whatever) that is functionally characterized" (Cummins, 1975, 741). The etiological solution consists in assuming that a trait exists as the result of previous selection processes: if the proper function of a trait type changes, that trait *type* may not be selected anymore and its presence in a population would decline. In organizational accounts, a trait exists because it performs its proper function. As this view of function is grounded in the individual level, if a trait *token* does not function properly, then that very same *token* may cease to exist. This is a consequence of the intimate connection between biological function and self-maintenance. A living system constantly struggles against its thermodynamic conditions, in such a way that the interruption of the regenerative process of its constitutive dimension entails the extinction of the system. If the teleofunction of a trait must be tied to its own existence, a theory of functions must therefore be able to explain how proper functioning entails the preservation of the

trait, while malfunction may imply the disappearance of the trait from nature. That is indeed what the theory of organizational functions suggests: proper functioning contributes to self-maintenance while malfunctioning does not. Therefore, the functionality of the organism is a prerequisite for its own existence: the interruption or modification (due to malfunction or severe environmental conditions) of functionality in minimal agents may entail death.

A distinctive character of organizational functions is their reliance on the current organization of a system undergoing physiological processes: "the organisational approach takes into consideration the relation between organism and environment as it unfolds in the present, in terms of internal compensations for environmental perturbations" (Mossio & Bich, 2017, 1098). So the temporal dimension is extremely short for organizational functions and it only ranges over the physiological process involved. In metabolic processes, the temporal scale is limited to the cycle of environmental input (of matter and energy) – internal processing (transforming and using the received input) – outcome (the production of waste discarded into the environment). In other cases —e.g. digestive systems, blood circulation, thermal regulation, and so on—- we also find cyclical processes that specify the time scale on functional ascriptions from an organizational point of view. So this is a major difference from etiological functions, which are not based on the current activity of the system.

As argued by Mossio et al. (2009), organizational functions occupy a midposition between etiological functions and Cummins-functions, keeping the advantages of each while avoiding their drawbacks. As they explain, organizational functions, like etiological functions, allow for the possibility of understanding functions in teleological terms, i.e. as natural teleofunctions. Unlike Cummins-functions, organizational functions are related to the *telos* of the organism. Moreover, organizational functions, like Cummins-functions, determine the function of the trait in relation to the current activity of the system. Unlike etiological functions that rely on the history of populations, organizational functions do not possess this historical dimension.

The temporal scale of organizational functions has both advantages and disadvantages. Earlier I enunciated the two *desiderata* that a theory of teleofunctions must meet. In connection with the Actuality *Desideratum* (AD), introduced in Section 5.2, I stressed that we should be capable of explaining the difference between goal-directed systems and non-goal-directed ones on the basis of their *actual intrinsic properties*. Organizational functions comply with this requirement. However, the Historical *Desideratum* (HD) presented in Section 3.2 states that functions must *have a historical dimension* to be treated in teleological terms. Such a historical dimension is not present in organizational functions. The physiological scale is not about how the interaction between the organism and the environment

became attuned during ontogenesis, or about how the constitutive dimension became stabilized. This dimension rather concerns an already adjusted and stabilized autonomous system and it just determines why it is teleological and why its parts must perform specific functions. I shall call *The Problem of the Missing Historical Dimension* this incapacity of autonomous systems theory to deal with the HD.

A related problem is that autonomous systems theory does not explain how autonomy is constructed. To be sure, an organizational account, by lacking a historical dimension, also lacks an explanation of how such organizational functions were constructed through time. This is a developmental question and a critical missing ingredient in this view. I shall call *The Problem of Construction* in autonomous systems theory this disconnection that exists in most contemporary accounts between how an autonomous system is analyzed in terms of its current organization with an analysis of how such an autonomous system came into being. As I expect to show, to account for this problem we should move away from a physiological scale and adopt a developmental perspective. So the Problem of Construction and the The Problem of the Missing Historical Dimension are two shortcomings of any purely physiological account. While this does not necessarily mean that these kinds of accounts are misguided, it certainly calls for a revision capable of somehow complement them with a developmental perspective in order to construct a consistent understanding of teleological agency at the individual level.

6.4.3 Normativity in autonomous systems

The account of normativity is connected to the account of natural purposes in minimal agency, i.e. to self-maintenance. As the main purpose of any minimal agent is the self-preservation of the system, a specific part of the system will be functioning correctly if it contributes to the self-maintenance of the whole system.

An account of biological normativity from the perspective of autonomous systems theory was developed by Christensen and Bickhard (2002) around the notion of *dynamical presupposition*. Let me remark first, that in autonomous systems the function of any system's trait is tied to its place in the whole functional unit. In any kind of process at the physiological time scale, different traits contribute to achieving a particular function. Normativity also rests on the interaction between the different parts of a system that allow the system to achieve a certain goal. The dynamical presupposition between the different parts of a system captures the idea that the normative judgments on traits are relative to the place of such a trait in the whole system. As expected, norms are related to goals: a system must do whatever it has to do to achieve a certain goal. As the primary goal in living systems is self-maintenance, the central idea of normativity is that a trait must do what is presupposed to do to self-maintain the system in viable conditions. As

noted, what a trait is presupposed to do is determined by the whole system: in physiological processes, different traits interact to produce a viable outcome. In this way, each trait presupposes the functioning of other traits and each trait is presupposed for the functioning of other traits. Consequently, a trait in a physiological process is "normative if it is dynamically presupposed by other processes in their contribution to the overall self-maintenance of an autonomous system" (Barandiaran & Moreno, 2008, 329).

In an interesting paper, based con Georges Canguilhem's ideas on normativity (see below), Barandiaran and Egbert (2014) distinguish between norm-establishing and norm-following. The former notion concerns the space of possible actions and their adaptive consequences that a system can perform according to its constitutive and interactive life conditions. They call this space the *normative field*. Within the normative field, there are two main regions: the *viability region* that refers to those activities (both at the constitutive and regulative levels) that self-maintain the system in viable conditions, and the *precarious* region that defines those states in which, under the same environmental conditions, the organism will die if it does change its activity. The normative field allows us to define how the system follows the norm established by evaluating, in normative terms, how a system acts according to a norm, i.e. how the system navigates within the normative field. So norm-following and norm-establishing, in relation to a normative field, allow us to recognize, model, and evaluate the system activity in physiological processes in relation to a norm.

The intrinsic dimension of normativity proposed by autonomous systems theory has among its main precursors the works by the French physician Georges Canguilhem on pathology (Canguilhem, 2012) and those of German-American neurologist and psychiatrist Kurt Goldstein on individuality (Goldstein, 1934). Both thinkers shared the idea that the norms of a living system are established by the system itself. Abnormal traits cannot therefore be defined on the basis of extrinsic parameters, but only from the point of view of the organism: "there is only one relevant norm; that which includes the total concrete individuality; that which takes the individual as its measure" (Goldstein, 1934, 269, quoted in Gayon, 1998, 310).

The account of normativity of autonomous system theory is non-standard; nor does it fit our folk use of normative talk in the life sciences; no wonder, since Canghuillem's and Goldstein's views were radically non-orthodox at their time. The autonomous systems view, coherently with its organicist foundations, rejects any supra-organismal account of normativity. Interestingly, however, supra-organismal views of normativity represent the standard attitude both in biology and in folk thinking. As noted, the SETF defines teleology, normativity, and proper function at the level of trait types, not trait tokens. So evolutionary, popu-

lational accounts evaluate tokens in terms of the types they belong to. Populations, not organisms, are the measure of normativity. In medicine, and in most clinical contexts, the well-functioning of a trait in physiological processes is typically assessed from a populational —but not necessarily evolutionary— perspective. In this case, albeit implicit, a statistical notion of normativity is at stake. Although it is clearly understood that traits contribute to maintaining the system healthy, standardization comes from an extrinsic, populational stance: a heart that pumps blood but does it differently from other hearts is not a normal heart. Also, our folk conception of normativity usually evaluates living beings —humans included— on the basis of generalizations and standardizations about how things ought to be. Someone, therefore, is labeled as abnormal if he or she does not fit such standardization. In our daily life, extrinsic normativity is the rule and results from a mix of biological, medical, socio-political, and religious views on normativity. In this sense, the principal understanding of normativity is from an extrinsic stance: "disease is a state only in relation to another state that has already been established as normal" (Keller, 2010, 45). As Keller insightful analysis demonstrates (cf. Section 5.2.3), while an intrinsic account of normativity is based on the analysis of individual traits, extrinsic and most popular (within and beyond biology) accounts state that normativity is an analysis of *trait differences in a population*: whether a trait is normal or not is determined by its difference with another, extrinsic parameter.

In contraposition to this, the autonomous systems view posits that norms are defined by each individual. As noted, a trait functions normally if it does what the system asks it to do. Abnormality thus can only be established in relation to such a norm: if a system constructs a certain norm about how a trait should work that is drastically different from the same trait in other individuals, such trait should be considered normal even if it looks abnormal from an extrinsic point of view. The fact that this opposes standard biological accounts should not come as a surprise. The view on teleological agency defended here might be a step forward towards a new evolutionary synthesis; cf. Chapter 8. Moreover, the rejection of extrinsic norms does not necessarily mean that extrinsic parameters are not seen as useful. It just means that natural norms are not extrinsic norms. That said, let me add that extrinsic norms —e.g. those derived by statistical analysis— may certainly be heuristically useful in biological research, they may have a practical value in medicine, and (perhaps) they may be useful in social interactions.

Only when this a priori counter-intuitive character of self-established norms is removed as a problem and perceived as the result of a long (and misguided) tradition on how normativity must be conceived, we become able to appreciate its strengths. The most significant one is that the autonomous systems account overcomes epiphenomenalist challenges to the SETF. There are two epiphenomenalist

challenges. In Section 5.1.2 I presented one of them. Norms defined in evolutionary terms are not causal but epiphenomenal, to the extent that evolutionary processes do not rest on populational causation but on individual causation. The second challenge was first articualted by Mark Bickhard and we already said something about it before. If norms are defined in terms of trait types, normativity is not analyzed from the perspective of what the organism does, but from the perspective of what populations have done during evolution. So, as I concluded in Section 5.1.2, selected-effect functions are epiphenomenal both at the individual level (as Bickhard suggested) and at the populational level (as I suggested, on the basis of the statisticalist view of natural selection). Yet norms in autonomous systems are not epiphenomena but they have a solid causal ground: the aforementioned $A-B-C-A$ chains present in the constitutive and interactive dimensions. The dynamic presupposition between parts of the organism and the whole organism in which norms rest identifies the causal loops that exist in physiological processes. The circular dynamics that produce and maintain the system define the norms, in such a way that the norms are intrinsic to the causal role that a part of the system has within the circular chains of causal interactions between parts and wholes.

The remarks in the previous paragraph run parallel to the aforementioned point I emphasized when discussing organizational functions. This concerns the Historical *Desideratum* that any account of teleofunctions must comply with. On pain of fostering a replicator, populational account, norms must rest on individual-level phenomena. Organizational functions comply with this *desideratum*, and this is the case with autonomous normativity insofar as the normative field is analyzed on the basis of the current organization of the system. However, as was the case with organizational functions, the *desideratum* is not observed in this account: autonomous norms have no historical dimension because they are deployed at the level of current activity of the system.

A historical perspective on autonomy is central to explaining how the autonomous system is constructed, with its organizational functions and with a particular normative field. So the shortcoming of not accounting for the Historical *Desideratum* is connected with another problematic issue, which I referred to as the Problem of Construction in autonomous systems. Explaining how norms are established on the basis of their current organization is not the same as explaining how they have been constructed. In other words, we can evaluate the functionality and normativity of a system by looking at its current organization, but this does not give us an explanation of how such an organization came into being, or of why it has the normative field that it has. This is a different, developmental question that autonomous systems theory scholars do not address. For the normative dimension, a constructive account is needed to understand how autonomy is achieved through an ontogeny that gives rise to a particular normative field.

6.5 Summary

I started this chapter by reviewing some significant categorizations for the analysis of natural teleology. Starting with Waddington's three temporal levels, this chapter focused the physiological one. Next, and considering the three possible strategies to confront the problem of teleology (Kant's Puzzle) in biology, this chapter discussed the naturalization of teleology at the physiological scale on the grounds of contemporary scientific accounts anchored in a rich historical tradition.

The principal precursor of the project of naturalization was Kant. Although he was not a naturalist, he nonetheless contributed two central insights that have proven critical for the later development of an agential view of teleology: the self-properties of living beings as the manifestation of natural purposes, and the causal loops —or $A-B-C-A$ chains— in physiological processes capable of explaining how self-properties emerge without the necessity to appeal to any kind of backward causation.

An important, well-defined, and robust tradition in the philosophy of biology emerged on the shoulders of Kant's legacy at the beginning of the 20$^{\text{th}}$ century: organicism. In the niche of the Theoretical Biology Club, organicism was proposed as a third, mid-term position between non-naturalist vitalism and reductive mechanicism. The three main pillars of organicism are (i) the centrality of the organism as the proper unit of analysis of biological processes —and the concomitant rejection of theories based on sub- and supra-organismal factors; (ii) organization as the central process to explain the complexity, order, and adaptiveness of living beings —supporting an order-from-disorder perspective of biological complexity; and finally (iii) the autonomy of biology —and therefore its irreducibility to physicochemical sciences.

Kant's legacy and the pillars of organicism have received a extensive scientific and theoretical treatment during the 20$^{\text{th}}$ century by different scholars working in different but related areas. Particularly relevant are the works on cybernetics by Humberto Maturana and Francisco Varela, and the work on systems biology by Ludwig von Bertalanffy, as well as the thought of Stuart Kauffuman and Robert Rosen.

Autonomous agency represents the contemporary account of agential teleology at the physiological time scale. Based on the recent work by Moreno and Mossio (2015) I have drawn the connections between closure, constitutiveness and self-constrains, and between openness, interactiveness and process. These distinctions made possible our definition of autonomous systems as self-determined systems brought about through different interactive processes acting on inner constraints to self-maintain the system in viable conditions. Next I presented three core properties related to agentivity in order to see how autonomous systems accomplish them. In addition, I also discussed how this theory provides us with the means to

trace the boundaries of autonomous agency.

The views on autonomous agency have promoted the possibility of naturalizing teleological, functional, and normative talk at the physiological scale. In contemporary accounts, both Kant's legacy and the organicist pillars are present. Natural teleology and its functional and normative dimensions reside in the self-properties of living beings. They can be naturalized by understanding how they are produced by $A-B-C-A$ chains; that is, by the recursive character of physiological process. Moreover, autonomous agency is clearly tied to the three pillars of organicism as revived by autonomous systems theory: (i) the organism is the central unit of analysis, (ii) (self-)organization (among other self-properties) is essential in the understanding of complexity and adaptiveness, which (iii) makes biology an autonomous science —i.e. living beings can only be understood by looking at living beings (organisms), not to non-living, sub-organismal parts. The principal purpose of biological processes at the physiological scale is that of self-maintenance in adaptive conditions. Organismal activity is directed at preserving the system alive, healthy, and, capable of sustaining its life at *the edge of chaos*. The function of traits is related to the organismal whole, in such a way that the place a trait occupies within the functionality of the organism determines how this trait must work. Consequently, the normative dimension is related to how different traits are presupposed by other traits to perform their proper function: a trait is said to be functioning wrongly whenever it does not do what it is supposed to do according to the needs of the whole organism and the dynamics and structure of physiological processes.

I showed how current theories on autonomous agents provide a consistent, scientifically robust, and well-developed account of teleology at the physiological scale. Surely, this does not mean that this account is complete and many interesting contributions are surely forthcoming in the near future. Notwithstanding, I foresee three problems that autonomous systems theory at the physiological scale is bound to face:

1. The problem of construction.
2. The missing historical dimension.
3. The problem of explanatory scope.

The first one —the *problem of construction in autonomous theory*— concerns the fact that autonomous systems theories, by being limited to the physiological scale, do not provide a full view of how autonomous systems are constructed. This is relevant insofar as the process of constructing an autonomous system is the

process responsible for building the normative and functional dimensions that autonomous systems theory analyzes at the physiological time scale.[16] Moreover, the second shortcoming of autonomous systems theory was formulated also in connection with the problem of construction, namely that autonomous theory does not comply with the Historical *Desideratum*. In the forthcoming chapters, I will argue that a developmental account of teleology is capable of addressing both issues: it is capable of explaining the construction of autonomy and it also incorporates a historical dimension on how functionality and normativity are established through ontogenetic processes. Finally, what I call *the problem of explanatory scope* of the physiological scale has to do with the fact that many important goal-directed phenomena remain unexplained just by focusing on the physiological scale. A developmental perspective will be central for understanding how complexity is created during ontogeny, how plastic responses take place, or how new variation arises, which are some of the central themes in contemporary biology. Thus, if teleological explanations are needed to explain the aptness of living beings, only adopting the physiological perspective of autonomous systems theory would inevitably have as a consequence the neglect of many phenomena that are crucial to explain precisely what we want to explain. A developmental perspective will help to recognize how major biological phenomena should be treated in adaptive terms as a consequence of being goal-directed.

Besides these three shortcomings of autonomous agency, even though my aim concerns primarily the developmental scale and not the physiological one, there are three important reasons why the theory of autonomous agents has been discussed here in detail:

1. A central aim in this thesis is to provide an analysis of natural teleology. Autonomous agency provides a view of one of the temporal scales (the physiological one) that needs to be taken into account in order to provide a complete analysis of natural teleology.

2. Autonomous agency captures an important trend of thought in the history of biology with significant advances that have given rise to new research disciplines. Moreover, autonomous agency represents a highly debated and central theme in the contemporary agenda in the philosophy of biology.

3. Last, but not least, autonomous agents, and the goal-directedness of physiological activity —from metabolism to behavior, will be an important piece in the proposal to be developed in the following chapters concerning teleological development. Many insights discussed in this chapter will play a key role in the exposition and defense of a teleological theory of development.

[16]The connection between autonomous agency and development has not been investigated yet; for example, Moreno and Mossio (2015) say nothing about it in their book.

Chapter 7

Teleological agency at the developmental scale

> The highest law of life, which connects all plan in time, has been named 'directedness' by K. E. v. Baer.
>
> Jakob von Uexküll, 1936, 144

> Natural forces which are not directed to an end cannot produce order.
>
> Karl Ernst von Baer, 1886b, 88

> Information is conceived to be a special kind of cause among all the factors that may be necessary for a phenomenon, the cause that imparts order and form to matter.
>
> Susan Oyama, 2000b, 3

7.1 The big question: teleological agency in development

In the previous chapter, I introduced an account of teleology based on insights taken from Kant's legacy and the organicist school. This view posits organisms as self-organized agents pursuing goals in relation to a self-established norm to self-maintain themselves in an adaptive state. This view naturalizes teleology at the physiological level. Unlike the account advanced by the MS, I see this approach in greater correspondence with current research in biology concerned with the organizational dynamics and the interactive processes of living beings. In this chapter, I will turn my attention to the developmental time scale. My

main aim is to sketch a theory of teleological development with as much detail as possible. This is the reason why this task is divided between this chapter and the following one. The function of this chapter is mainly expository, preparing the ground for Chapter 8, where I shall bring together all the threads in order to set up my own proposal. Thus, here I will first put forth a number of considerations concerning teleological development and next I will also introduce some of the main forerunners of teleological development in the history of theoretical biology. With these preliminaries in place, I shall eventually turn to consider the contemporary view on developmental teleology. I will conclude the chapter with a summary of the conceptualization and philosophical work that in my opinion is still needed in order to provide a robust account of teleological development.

The chapter is organized as follows. In this section, I identify what in my opinion are the two most relevant explanatory aims of a teleological theory of development. Section 7.2 provides some relevant historical background and introduces a number of key points in the thought of some important precursors of teleological notions in development. In Section 7.3 I undertake a revision of the literature on contemporary attempts to understand teleology in development, paying special attention to the work of Denis Walsh on the *Agential Perspective*. Towards the end of Section 7.3, I will identify an important missing element in Walsh's account, while singling out at the same time the necessary ingredients to fill this gap. This is the task of Chapter 8.

It is unquestionably the case that no other organic process conveys as strongly as development the feeling of goal-directedness: a process starting from a single cell and going through different stages to construct an adaptive adult organism. Each step is determinant for the adaptive value of the next one. The Estonian embryologist Karl Ernst von Baer will be one of the main historical characters in this chapter, as he often used the example of the metamorphosis of butterflies to call our attention to how the different ontogenetic stages going from egg to adult, through larva and pupa, are necessary for the future development of the butterfly. "How is it possible to mistake that all of these operations are ordered with respect to a future need? They are directed to that which is to come into being", wrote von Baer in one of his posthumously published papers (von Baer, 1886b, 58). Development may therefore be characterized as a chain of complex processes that results in a highly functional and well adapted to its environment organism, where every element in the chain is necessary for the occurrence of the next one. But, why is it so important to treat development in teleological terms?

There are two main explanatory aims that *any* teleological theory of development should pursue, which give us the two basic reasons why an analysis of teleology at the developmental scale is crucial:

(i) the connections between development and evolutionary theory, and

(ii) the limits of the physiological account I discussed at the end of Chapter 6.

Let me deal with each point in detail.

Development and adaptive evolution. Point (i) above has to do with recent attempts at building a new or modified evolutionary synthesis that the supporters of Developmental Turn (DT) have been pursuing in the last three decades. Recall that the hypothesis that development comes first goes together with three basic assumptions, namely (a) that *phenotypic variation* is a consequence of developmental causes, (b) that *inheritance* is not about replication but about the reconstruction of traits during ontogenesis in each generation, and finally (c) that *fitness* is determined by adaptive developmental processes. Essentially, this is tantamount to claiming that the three core ingredients of natural selection occur at the developmental scale. Clearly, here there is also a connection with the statisticalist view of natural selection according to which adaptive evolutionary causes lie at the individual level and the main force in evolution is internal to the organism and not external to it. Hence, if development is going be at the core of our understanding of adaptive evolution, then we are compelled to inquire into the question whether the adaptiveness of development needs to be explained in teleological terms. This must be necessarily so inasmuch as our goal is the naturalization of teleology: if teleology is essential for explaining the aptness of living beings, and development is the process that brings about adaptive organisms, then, if we eschew the eliminativist stance, development must somehow be treated in teleological terms. This motivation is also connected to one of the three problems of the physiological scale discussed in the previous chapter, namely the *problem of explanatory scope*. In a nutshell, as the account of teleology at the physiological scale leaves many adaptive processes out of its scope of analysis, the developmental scale is needed to deal with them and, concomitantly, to connect them with evolutionary processes.

Complementing the physiological scale. Point (ii) concerns the fact that the physiological time-scale is limited to the analysis of the current organization of the system. This motivated our pinpointing of the other two problems faced by all analyses circumscribed only to the physiological scale: the problem of construction and the missing historical dimension. As soon as we confine teleology to the physiological time scale, we directly miss the possibility of offering an account of how traits are constructed, how adaptive variations come into being during development, and how the functionality and normativity of agents are achieved —not just maintained. All these processes appear to have a teleological, goal-directed flavor. Even though the relationship between the physiological and the

developmental scales is a fairly close one, one thing is to see the actual organization in teleological terms, and another, very different issue is to come to know how such organization was constructed during epigenesis.

As I see it, these are the two main reasons why a naturalistic analysis of teleological development is indispensable in current biological theory. Unfortunately, no fully worked out account of teleological development exists yet. This is an important difference with the physiological time scale. In spite of there existing eminent forerunners in the history of biology and a number of incipient contemporary proposals for introducing a teleological dimension in development, it is clear that current biological theory has not addressed this issue with the same emphasis as in the case of teleological physiology.

I suspect that the main reason for this shortcoming has to do with the fact that development has proven to be one of the toughest domains to naturalize teleology. So, before moving on, let me provide a brief sketch of why this should be so. Unlike the physiological and evolutionary time scales, there are no $A - B - C - A$ chains in development. That is, the adult phenotypic stage does not seem to be providing any kind of circular feedback to previous ontogenetic stages. At the physiological time scale, circular causal chains are possible due to the relation existing between parts and whole which is mediated by inner regulations at the constitutive and interactive level. At the evolutionary scale, inheritance establishes the circular relation between generations mediated by selection processes. Such circularity is not present in development. The arrow of time points in one direction only, it does not bend or turn backwards. As a consequence of this, the main strategy applied to naturalize teleology —adaptive biases acting on $A-B-C-A$ chains— seems not to be available at the developmental time scale. This imposes severe constraints over any teleological theory of development. But, and this qualification is important, the absence of $A - B - C - A$ chains in development is a direct consequence of assuming that adult states are the goals of development. In the next section, I will argue this is a mistake and that adult states cannot be taken to be the goals of development.

7.2 Forerunners

In this historically-minded section, I shall review the work of some important precursors of the contemporary standpoint favoring teleological explanations in development. At the end of each subsection, I will summarize and systematize what in my opinion are the main take-home messages that we should keep in mind for further analysis. I think that there are three main landmarks in the history of biology that will be of particular interest to us: the preformation–epigenesis debate, the teleomechanicist program, and as it was also the case for

the physiological scale, the organicist tradition.

7.2.1 The preformation–epigenesis debate

The preformation–epigenesis debate took place principally in the 17th century and the first half of the 18th, and it will of great help to foreground three points that will prove central in our journey towards a view of teleological development aligned with the DT.

Information: the cause we need

The old debate revolved around morphology. Is morphology present at the time of conception or is it acquired through a process of piecemeal construction? The former view was defended by preformationists, while epigeneticists supported the latter. For any contemporary observer, however, it should intuitively be quite clear that at the time of conception there is no fully formed organism with all the morphological features of the adult organism, that ontogeny is not just growth of an already structured organism. Quite the contrary, morphology comes into being through a step by step process of construction traditionally referred to as morphogenesis, i.e. through epigenesis. Development is not just the growth of parts but the construction of a complex organic structure.

It would nonetheless be debatable to claim that the epigenetic team won the battle. To be sure, perhaps epigeneticists won the contest in the 17th century, but a new kind of preformationism emerged in the 20th century, namely genetic preformationism. The idea was already presented in Chapter 2 and discussed in Chapter 4. According to the MS view —from the work of Weismann to the popularization of the Central Dogma and Dawkins' gene-eye view— genes carry the necessary information to produce an apt trait. Surely, no morphology is present in a fertilized egg. What is present is the genetic information needed to produce such a morphology. Many metaphors, such as developmental programs, genetic instructions, or genes-for, are representative of the attempt to stress that the information needed for development does not come from epigenesis but is already preformed at the time of conception.

The preformation–epigenesis debate is connected with two alternative ways to understand how order is created in living beings: order-from-order and order-from-disorder. The former represents a preformationist viewpoint, claiming that order pre-exists development. Order comes from already ordered units of inheritance endowed with all the information needed for producing a trait. As Oyama (2000b, 2) states, the core of "a 'preformationist' attitude toward information" is that "it exists before its utilization or expression". In this sense, as the MS assumes, genetically inherited information precedes its expression. Preformationism persists

in biology after all. In contradistinction, the order-from-disorder strategy is representative of the epigenetic camp: order originates in development, development organizes disordered matter.

The discussion regarding whether order is preformed or arises during epigenesis is usually presented in informational terms. This is so because, as Oyama eloquently explained, the principal role of information in biology is to be "the cause that imparts order and form to matter" (Oyama, 2000b, 3); information is seen as "the modern source of form" (Oyama, 2000b, 1), as the responsible of causing such complex and well-organized systems found in the living realm.[1] The first point of this subsection is thus the idea that information is the key concept in the search for a teleological view of development: the question about teleology is a question about biological information. The connection between teleology and information is quite clear. If teleological explanations are needed for explaining the complexity and order of adaptive systems, and information is such a cause that imparts order, then, explaining information in biological systems provides the key to natural teleology. Note that Oyama defines information as a *cause*. The naturalization of teleology is therefore about identifying the causes of teleological explanations. Explaining development in informational terms is thus a path toward understanding the causal basis of teleological development. Chapter 8 is an attempt to walk this path.

From informational preformationism to the ontogeny of information

In Section 4.1.1 and Section 4.2 I distinguished between two kinds of questions at play: material questions and conceptual questions. Material questions were about the material nature of genes and systems of inheritance. I argued that the molecular gene should be replaced by the reactive genome and that the idea of genes as the unique source of inheritance should be complemented with extended systems of inheritance. However, the most relevant questions about genes and inheritance are conceptual: one about the status of genes as units of development, and the other about inheritance as a process of replication.

Concerning the first question, I presented in Section 4.1.1 the notion of the *ontogeny of information*, introduced by Oyama (2000b), to challenge the idea that information in development is preformed, to be replaced by the idea that the cause that imparts order to development (i.e. information) comes from development itself. Regarding the second question, I supported in Section 4.2.2 a developmental conception of inheritance in opposition to a replicator stance. Once we move to a developmental view of inheritance no preformationism is possible: inherited information cannot precede development because inheritance is part of development.

[1] See also Moreno (1998, 203): "The type of causal action of information is 'formal' in the sense that it infuses forms, i.e., it materially restructures matter according to a form".

Modern Synthesis	Developmental Turn
Order-from-order	Order-from-disorder
Genes as the unit of development	The organism as the unit of development
Replicator conception of inheritance	Developmental conception of inheritance

Table 7.1: Alternative ways of explaining order: part II.

No order in inherited units preexists development insofar as inheritance is about the very processes of constructing a trait (i.e. development).

The conclusion of the aforementioned conceptual question is that the order-from-disorder view should be supported in light of how development is conceived of by the DT. This analysis was summarized in Table 4.3 on page 111, reproduced here as Table 7.1 for convenience.

With this analysis in mind, we can turn to the preformation–epigenesis debate. Here, it is also relevant to distinguish between two kinds of questions:

1. Material question: Are traits preformed in the DNA?
    - From morphological preformationism to morphogenesis.
    - From genetic preformationism to reactive genomes.
2. Conceptual question: Does order precede development?
    - From informational preformationism to the ontogeny of information.

We know that once we look at reactive genomes in the post-genomic era (Section 4.1.1), traits cannot preformed in DNA. So, concerning the material question, the answer is negative: "everything needed for producing a phenotype is not prepackaged in the fertilized egg" (Gilbert, 2012, 20). However, as it was also the case in our discussion of genes and inheritance, the material and conceptual questions about preformationism are independent: changing the answer to the material question does not necessarily entail changing the answer to the conceptual question. That this is so is made clear by taking a look at the history of biology. Even if morphological preformationism was wrong, another piece of matter may be there to support preformationism: DNA. The fact that the mainstream answer to the material question deserves revision does not entail that the conceptual question also needs to be challenged. This is why the material and conceptual questions are distinct. Accordingly, the conceptual question concerns the possibility that reactive genomes also allow for a sort of preformationism. It is certainly possible

Modern Synthesis	Developmental Turn
Order-from-order	Order-from-disorder
Genes as the units of development	The organism as the unit of development
Replicator conception of inheritance	Developmental conception of inheritance
Informational preformationism	The ontogeny of information

Table 7.2: Alternative ways of explaining order: part III.

to support this view: information for developing traits is not present just in DNA but is distributed among different genetic and non-genetic systems of inheritance. Traits are preformed in distributed developmental resources. This is tantamount to supporting an extended replicator view.

In this sense, to challenge preformationism is not just to say that genes do not code for traits. What we have to do here is to argue that information is created in development: that the specification of developmental outcomes resides neither in genes nor in non-genetic resources. This is the idea of the ontogeny of information. So the search for a non-preformationist view of development is to understand that development is not about unfolding inherited information, but about the creation of information during development. My second point in this section is that *the ontogeny of information, together with a developmental conception of inheritance, is central for moving beyond any preformationist stance.* Table 7.2 extends Table 7.1 on the preceding page highlighting alternative ways —preformation or epigenesis— to understand biological information in connection with different interrelated issues discussed throughout this thesis. This is crucial if we take into account that informational preformationism was the key to naturalizing teleology within the MS framework. Evolved information imparts order in the living realm, so the naturalization of teleology rests on the evolution of information. But the search for a view of natural teleology beyond the MS requires an epigenetic view of information; a view where information is in-formation during development, not pre-formation by evolution.

Goals cannot be about adult states

The emphasis on a developmental theory of inheritance and the ontogeny of information leads us to the third main point of this chapter: *developmental goals cannot be about adult states.* The structure of the argument to defend this idea is the following:

1. Adults states cannot be preformed (by the ontogeny of information).

2. Having a goal must be a causal factor in development (a main prerequisite for naturalizing teleology).

3. Therefore, developmental goals cannot be about adult states.

Let us assume, for the sake of the argument, the negation of my conclusion, namely that developmental goals *are* about adult phenotypic states. If this is so then either point 1 or point 2 must be false. In other words, if we support point 2 then point 1 is false (let's call this option a) and if we support point 1 then point 2 is false (option b):

(a) On the one hand, one might claim that goals are about adult states and that they causally contribute to development —in support of point 2. This position is, for instance, the one defended by Mayr's teleonomy: DNA sequences causally unfold traits. But in this case, goals would be (genetically) preformed; consequently, against point 1, the information needed for development preexists developmental processes.

(b) On the other hand, one might claim that goals are about states and that they are not preformed —in support of point 1. However, in this scenario, how does having a goal can possibly contribute to development? If the goal is about a future stage and such future stage is not present in any way in any particular developmental stage, then goals cannot causally contribute to development unless some sort of backward causation is posited, an idea that goes against point 2. This scenario is relatively similar to the one pictured by vitalists. Most of them were embryologists and their reason for positing vital forces was their powerlessness at the time of understanding how development could be directed to adult states. So perceiving directedness toward adult states is what lead them to a non-naturalist position (i.e. against 2), as I argued here.

Thus, to conclude so far, if goals were adult phenotypic states, then either we should support some sort of preformationism, or we should posit some sort of backward causation. I conclude, therefore, that the goal of development is not oriented toward an adult state.

To summarize, I wish to keep the following points from this subsection:

1. The translation of the question about teleology to a question about information in biology.

2. *A developmental theory of inheritance* and the notion of *the ontogeny of information* are central to avoiding any sort of preformationism.

3. *Developmental goals are not about adult states*: If goals were about adult states, then, or adult states would be preformed (against the view of the ontogeny of information), or having a goal would not causally contribute to development (against naturalization).

7.2.2 Teleomechanicism

Teleomechanicism is the label proposed by Lenoir (1989) to refer to the post-Kantian German biology that saw its heyday between the end of the 18th century and a large part of the 19th century, with a huge amount of empirical advances in embryology and morphology, and a defined philosophy of biology. Different biologists since Kant contributed to the gestation of teleomechanicism, and the influence of Kant is clear (but see Zammito (2012) for some disagrement). Kant's work was not only an important influence on the physiological scale but also on the developmental one. Indeed, it is relevant to note that the distinction between the physiological and developmental scales was presented here in order to systematize the analysis of natural teleology, even though I recognize that this distinction is not so transparent in the history of theoretical biology. As we will see, many important figures used physiological phenomena to understand teleological development and vice versa.

Be that as it may, teleomechanicism inherited from Kant exactly that which the word itself denotes: a commitment with both mechanistic and teleological explanations to understand the living realm. On the one hand, teleomechanicism recognized that its main research areas —embryology and morphology— study the different mechanisms of development and that such mechanisms obey the laws of physics and chemistry. On the other hand, like Kant, teleomechanicists also accepted that a teleological understanding of living beings is inevitable; i.e. that a complete comprehension of living beings requires the introduction of teleological language. The main footprint of Kant is the commitment to a regulative view of teleological explanations. Most scholars within teleomechanicism adopted Kant's picture of mechanistic and teleological explanations to understand the discoveries in their research areas (Lenoir, 1980).

In this subsection, I will only focus on one the main characters of teleomechanicism: Karl Ernst von Baer. There are six reasons why von Baer deserves special attention and why his view will be important in this chapter. First, von Baer was the main architect of teleomechanicism (Lenoir, 1989, 16). Lenoir pictures teleomechanicism as a progressive program, where each actor "criticizes, refutes some aspect of, and ultimately extends the explanatory domain of its predecessor" (Lenoir, 1989, 13), with von Baer's being the most advanced view within teleomechanicism. Second, he could be considered the founder of embryology, one of the central areas of developmental biology. Third, he proffered a number of penetrat-

ing reflections on natural teleology and development, especially in von Baer (1886a, 1886b). Fourth, he considered teleology not just as a regulative, heuristic tool for biologists, but as a real, immanent phenomenon intrinsic to the dynamics of development. Fifth, and in connection with the previous point, he did believe that teleological explanations could be legitimately included in science. And, finally, his proposals on teleology were discussed in connection with Darwin's theory, so he provided an alternative view of living aptness that stood in explicit contradiction with the theory of natural selection. In this sense, while inheriting many views from Kant, von Baer moved away from Kant's insight on natural teleology.

Ziel und Zweck

The first difference with Kant's view is terminological. Von Baer insisted on the importance of contrasting purposes and goals. This distinction, as Lenoir (1989, 272) points out, is difficult to grasp in "Western languages". This difficulty concerns the subtle semantic nuance between the German words *Zweck* (roughly translatable as 'purpose') and *Ziel* (typically rendered in English as 'goal'). The motivation for this distinction is directed at finding a legitimate place for teleology in the natural sciences. *Zweck* has clear volitional connotations and the traditional neglect of teleology within biology comes in part from perceiving an association of purposes with intelligent design (such as human intentions or God's wishes). But for von Baer, this distinction was principally introduced to discuss Darwin's view on natural design. Consider the following words by von Baer:

> Nearly a century ago Kant taught that in an organism all the parts must be viewed as both ends and means [Zweck und Mittel] at the same time. We would rather say: goals and means [Ziele und Mittel]. Now it is announced loudly and confidently: Ends [Zwecke] do not exist in nature, there are in it only necessities [Notwendigkeiten]; and it is not even recognized that precisely these necessities are the means [Mittel] for reaching certain goals [Ziele]. Becoming [ein Werden] without a goal [Ziele] is simply unintelligible. (von Baer, 1886b, 231; quoted in Lenoir, 1989, 271)

Let's analyze this quote. First, when von Baer writes "now" [jetzt] he is referring to the time in which Darwin's theory (and also Haeckel's views) was presented. So Darwinians are identified here as those who deny the existence of any kind of purposefulness. Rather, as natural selection theory maintains, the aptness of living beings is explained by chance variations and their necessary consequences in phenotypes. To recapitulate from Chapter 2, as Monod (1971) famously defended, chance refers to the randomness of variation, while necessities concern the material consequences that such hazardous variants cause. As already explained,

this led evolutionary biologists to abstain from looking at development as part of evolutionary theory. The point that von Baer is trying to make is that the very necessities stressed by Darwinians are the result of the goal-directedness of development. While Darwinians may dispose of any talk about purposes [Zwecke], goals [Ziele] must be part of the explanation of aptness: chancy variations without a goal do not produce apt traits, and the consequences of variations must incorporate the goal-directedness of development. Rather than looking for the necessary consequence of hazardous variations, Darwinians should investigate the *necessities of developing systems* in the production of apt traits. These necessities are embodied in Darwin's emphasis on the struggle for life, i.e. those individual processes that define the individual fitness of an organism. Von Baer defended that without goals one cannot explain how aptness and other central biological phenomena take place. While he was not an anti-evolutionists, von Baer insisted on the idea that Darwin's theory is incomplete on this point, that we must recognize "that precisely these necessities [individual processes] are the means for reaching certain goals [apt phenotypes]". Getting rid of directedness leads to the misunderstanding of development: ontogeny without goals "is simply unintelligible". As Lenoir (1989, 271) points out,

> if the only principles admissible in science are those deriving from mechanistic necessity, then the most fundamental questions of zoology, namely those concerned with organization, generation, development, function, and adaptation, must remain ultimately unintelligible; for they must reduce to an accidental concatenation of mechanical processes without a common ground for their necessary interconnection. Only a teleological framework could serve as a corrective to this problem.

Therefore, the distinction between *Zweck* and *Ziel* was meant to avoid attaching any negative connotations to teleological explanations and to divorce teleology from anthropomorphic or religious views. But most importantly, the distinction was also intended to argue that Darwin's view rejected the existence of *Zwecke* but not necessarily the existence of *Ziele*. Indeed, *Ziele* are central to understanding how living things come into being —i.e. ontogenesis. The defense proposed by von Baer thus departs from Kant's emphasis on teleology as a mere regulative principle to take it as an intrinsic necessity for development to be possible.

Harmony

Moving away from *Zwecke* but retaining *Ziele* van Baer's is a clear attempt to legitimate the place of teleology —as a relation of means and goals (*Ziele*)— in

the explanation of adaptive development, but what naturalistic theory did von Baer propose?. Certainly not a very well-developed one —which would have been anyway quite difficult given the state of research in developmental biology at the time. We can nevertheless identify two fundamental ideas of von Baer's view that will prove central in this chapter: regulation and harmony.

According to von Baer, "every organism in the process of coming into being has a goal. Without goals how could anything subject to *regulation* come about?" (von Baer, 1886b, 180; emphasis in the original; quoted in Lenoir, 1989, 272). However, he certainly had no clue about how developmental regulation towards a goal could be understood in scientific terms. Fully aware of this shortcoming, von Baer resorted to a metaphor: musical harmony. According to this metaphor, the process of development consists of the *orchestration* of the material basis of development towards a goal to produce a system internally (in relation to the parts of the system) and externally (in relation to the environment) harmonized. Harmonization is therefore a relation between two elements —between parts of the system or between the system and the environment— directed towards a goal. As he explains,

> the reciprocal interconnections of organisms with one another and their relationship to the universal materials that offer them the means for sustaining life, is what has been called the harmony of nature, that is a relationship of mutual regulation. (von Baer, 1886b, 228–229.; quoted in Lenoir, 1989, 275)

The idea behind this metaphor is that harmonic regulation of interacting elements towards a goal produces adaptive phenotypic results, "for it is precisely the recognition of the mutual interrelation of all the processes in nature and the harmony in their institutions that gives them the greatest pleasure" (von Baer, 1886a, 51). Interacting matter with no directedness cannot produce such results. As von Baer eloquently stated, "only natural forces which are not directed to a goal can produce nothing regulated" (von Baer, 1886a, 88). In his view, the distinctive character of biology is the presence of harmonic regulation in development that leads us to appreciate the goals of living beings in their adaptive outcomes. What he lacks, but contemporary developmental biology supplies, is an account of how harmonic regulations occur in real developing systems.

To conclude, von Baer aimed to take up "teleology without regrets" (Lenoir, 1981): to assume that teleological explanations are intrinsic, immanent, irreducible, and, consequently, indispensable for understanding development. The possibility of the naturalization of teleology —even though naturalism is a 20$^{\text{th}}$ century construct— was within the scope of von Baer's theory. Goals are central for development to take place. Without goals, order is not possible, he said. The

"strategy of life" (Lenoir, 1989) is the use of matter to produce an organized being by goal-directed means. In our journey towards naturalization, I would like to hold onto the following three ideas from von Baer's work and his emphasis on *Ziele*:

1. *Natural forces which are not directed to an end cannot produce order* (von Baer, 1886a, 88). Bare mechanicism is not capable of accounting for organic development and the creation of order and complexity we find in living systems.

2. A central element in development is *regulation*: goals are achieved by regulating means, i.e. by regulating the dynamics of development toward a goal. Moreover, while von Baer could not provide a scientific view of regulation, he suggested a vivid metaphor: *musical harmony*. Goals are achieved by harmonic regulations during development.

3. Developmental regulation towards a goal is achieved through harmonic interaction *between the parts* of the developmental system and *with the environment*.

We saw in the previous chapter how Kant's insights were naturalized thanks to the advances in science dealing with physiological processes. In this chapter, I would like to suggest that the same scenario applies to development: recent advances in developmental science make it possible for us to comprehend developmental regulation from a scientific standpoint.

7.2.3 The footprints of organicism

Many adherents of organicism also worked in development. In this section, I shall introduce the work of four important thinkers within this school of thought, Jakob von Uexküll, Edward Stuart Russell, Ludwing von Bertalanffy, and Conrad Hal Waddington, who also addressed the problem of teleology in development and, as I shall note in the next section, who have had an influence on contemporary views.

Jakob von Uexküll

Jakob von Uexküll (1864–1944) was part of the organicist movement during the 20$^{\text{th}}$ century (Kull, 2001). His relevance here is also central, mainly because von Uexküll's thought shows a strong influence of von Baer's. The main contributions of von Uexküll to the aims of this chapter are twofold: the revival of von Baer's views and his introduction of the notion of *Umwelt*.

Concerning the connections with von Baer, let's note first that von Uexküll adopted von Baer's terminological choice: "men have spoken of 'purpose' and 'purposefulness' in Nature as a sort of human being [...] it is advisable therefore to

dismiss from biology, for all time, expressions such as 'purpose' and 'purposefulness' " (von Uexküll, 1926, 270). Instead of purpose, as Kull (2001, 6) notes, von Baer used the word "Zielstrebigkeit" which is "the term of Karl Ernst von Baer, from whom, clearly, Uexküll has inherited an understanding of the fundamental importance of the temporal organization of organisms". Another connection with von Baer is that, although von Uexküll was a Kantian, his view of teleology was not the regulative view proposed by Kant, but rather he appears to have been a realist concerning teleology: "while some teleo-mechanists treated organismic teleology more as a methodological commitment, Uexküll's use of romanticism instead of Kant's *Critique of Judgment* leads him to treat it as a fully real part of nature" (Feiten, 2020, 3).

Moreover, like von Baer, he also made use of a musical metaphor to explain the directiveness of development and the interactive regulation both with the environment and its inner organization (respectively, the *Umwelt* and the *Innenwelt*; see below):

> planned embryonic development [...] begins [...] with the three beats of a simple melody: morula, blastula, and gastrula. Then, as we know, the development of the buds of the organs begins, which is fixed in advance for every animal species. This proves to us that the sequence of formal development has a musical score which, if not sensorily recognizable, still determines the world of the senses. This score also controls the spatial and temporal extension of its cell material, just as it controls its properties. (von Uexküll, 2010, 159-160)

The second and most important point is that von Uexküll introduced the notion of *Umwelt* to refer to the subjective experience of the environment that each organism has, along with the notion of *Innenwelt*, to denote the experience of the inner environment of an organism. I will use the abbreviation U–I to refer to the Umwelt and the Innenwelt. This is clearly connected to the notion of agency, insofar as it requires taking organisms as capable of perceiving, sensing, and constructing their environment. The environment of the organism cannot be defined independently of how the organism experiences it. As von Uexküll (1923, 266) states, "nobody is a product of their environment — everybody is the master of one's Umwelt" (quoted in Kull, 2001, 1).

However, based on the notions of U–I, von Uexküll moves away from von Baer's view. As Ostachuk (2020, 164) notes, "[a]ccording to von Baer, an embryo possesses this 'effort toward a goal.' Uexküll does not agree with this argument. In the first place, Uexküll considers that the goal is not the adult organism but the congruity with its *Umwelt*". Thus, von Uexküll provides a key ingredient to overcome an important hindrance mentioned in the previous subsection, namely,

that goals cannot be preformed. In his view, ontogeny is not directed towards producing a particular outcome in the future; rather, development is guided by a constant response to the U–I, in a way that the embryo "unerringly produces definite counter-properties, which fit into a definite group of properties in the external world" (von Uexküll, 1926, 317). As von Uexküll stresses, the fit between organism and environment results from the adaptive regulation of developmental processes according to the environmental context of development. This is the strategy I would like to follow in Chapter 8. If development is adaptive due to the harmonic regulation between parts, this is so because developmental goals are directed toward the regulation of the U–I.

Edward Stuart Russell

E. S. Russell, together with von Bertalanffy and Waddington, contributed to the recognition of important elements of development that today help us to understand it as a goal-directed phenomenon. There are two interesting elements in Russell's attempt to understand goal-directedness in living beings: a terminological point that, once again, connects him with von Baer, and an explicit characterization of goal-directed activity.

First, as pointed out by Esposito (2013), there is a direct influence of von Baer's work in Russell's theory. Russell noted, in the same spirit as von Baer, that "one is tempted to use the word 'purposive' in the description of these activities, but this term is used in many senses and has a strong psychological flavor about it" (E. S. Russell, 1934, 836). So he adopted instead the term *directivness*, which is applicable to both psychological and non-psychological processes. That is, "the directiveness of vital processes is shown equally well in the development of the embryo as in our own conscious behaviour" (E. S. Russell, 1934, 836). This worries about terminological decisions are in part due to the negative connotations of teleology in biology: it is not so easy to accept that plants or bacteria have goals, purposes, or norms. That is why many naturalists insist on the importance of terminology (as it was also the case with Mayr's use of 'teleonomy'). In this case, Russell's aim is to emphasize the fact that no psychology or consciousness is needed. Organisms need not be aware of anything. Organic activity "can rarely be called purposive activity for the organic agent concerned is seldom explicitly aware of the goal towards which its action is directed, much less of the biological end which it subserves" (E. S. Russell, 1945, 3).

Besides the terminological issue, E. S. Russell (1945, 110) explicitly promotes the following five properties of goal-directive activity:

1. When the goal is reached, action ceases; the goal is normally a terminus of action.

2. If the goal is not reached, action usually persists.

3. Such action may be varied:

    - if the goal is not reached by one method, other methods may be employed;
    - where the goal is normally reached by a combination of methods, deficiency of one method may be compensated for by increased use of other methods.

4. The same goal may be reached in different ways, and from different beginnings; the end-state is more constant than the method of reaching it.

5. Goal-directed activity is limited by conditions, but is not determined by them.

The first two points are clear marks of goal-directed activity: if my goal is to write a thesis, I will keep working until it is written, and clearly, I will stop writing once it is finished. These two points deserve special attention. I will come back to them by the end of this chapter. For the time being, let me focus on the other three points. The third point states that there are alternative paths toward the same end. If a developmental pathway is not successful or possible, alternative pathways may be found. Similarly, developing organisms are capable of accommodating the trajectories in the presence of difficulties, be that due to changes in external conditions or due to changes in internal ones.

The different modes that organisms can find to achieve a goal are connected with the fourth point. It emphasizes a central element of organic activity, namely the *creativity* of development. *Creativity* endows organisms with the capacity of searching different paths toward the fulfillment of a certain goal. Although it is used in a metaphoric sense, the creativity of development concerns the fact that organisms constantly manage to find suitable outcomes for their conditions of existence. As E. S. Russell (1945, 144; emphasis in the original) expressed it: "what *is* distinctive is the active persistence of directive activity towards its goal, the use of alternative means towards the same end, the achievement of results in the face of difficulties", in a way that "the end-state is more constant than the method of reaching it". Creativity is thus connected with the issue of preformationism: if developmental pathways are creative —i.e. sensitive to the context of development— they are not pre-established or determined by previous development.

The creativity of development eventually takes us to the last property. As Russell claimed, "goal-directed activity is no mere resultant of material conditions, as is the case with inorganic systems [...] It is not dominated by conditions, but strives to surmount or utilise them in its movement towards its goal" (E. S. Russell,

1945, 144). This means that the material basis of development operates as a constraint that restricts the possible phenotypic variations but also allows for a repertoire of possible developmental outcomes. By saying that development "is limited by conditions", therefore, Russell means that the material conditions of development make variation finite, while at the same time such material basis does "not determine" the outcomes but defines a repertoire of possible outcomes. The search for suitable developmental pathways within the repertoire of possible outcomes is possible inasmuch as outcomes are not prescribed, but rather they result from a creative developmental process.

Ludwing von Bertalanffy

Like Russell, Ludwig von Bertalanffy also deals with an important phenomenon related to the teleological character of development, namely *equifinality* (von Bertalanffy, 1952, 1969). His view is nowadays well-known and captured under the idea of the *robustness* of development (Bateson & Gluckman, 2011), i.e. the fact that the same phenotypic outcome can be achieved by different developmental means. In other words, different organisms with different genetic underpinnings can build the same phenotypic outcome through different developmental pathways. As von Bertalanffy appreciated, this entails a sort of directedness of development, insofar as organisms manage to achieve a similar end by different means. There are two important remarks to be made in this connection.

First, von Bertalanffy's view of equifinality is connected with the openness of living beings: "In a closed system, the final state is unequivocally determined by the initial conditions", however, "in open systems, the same final state may be reached from different initial conditions and in different ways" (von Bertalanffy, 1969, 40). In this sense, the capacity of adopting plastic developmental pathways to achieve a robust outcome is related to the capacity of organisms to modulate their interaction with the environment. This idea will be developed later in detail.

The second point is primarily historical. The equifiniality of development was Hans Driesch's main motivation to adopt a vitalist position. In his experiments, he manipulated the egg in different ways and obtained the same developmental result. The fact that different initial conditions led to the same developmental outcome was, for him, evidence that a vital force was operating in the egg that pushed it toward a specific end, and that consequently "All believers in epigenesis are Vitalists" (Driesch, 1914, 39, quoted in Oyama, 2010. However, as it will be developed later on, the understanding of development forged during the 20th century, and principally during the 21st century, allows us to see that no vital forces need to be posited to understand equifinality.

Conrad Hal Waddington

Conrad Waddington also appreciated the phenomenon of equifinality by stressing the importance of canalization in evolutionary theory. His theory, restored and refurbished later by West-Eberhard (2003) (cf. Section 4.1 for details), suggests that canalization allows to anchor phenotypic outcomes in reliable inherited materials. In other words, while phenotypic variation may arise from changes in developmental resources, the developmental system manages to canalize such variation into more robust developmental pathways capable of being reliably inherited. Canalization, therefore, also illustrates the robustness of developmental pathways.

However, another phenomenon that also exemplifies the adaptive character of ontogeny, also studied by Waddington, is that of phenotypic plasticity. As it was explained in Section 4.1, plasticity was reduced to the genetic level by the MS in order to make it tractable within its framework. However, once phenotypic plasticity is rather seen as an organismic phenomenon taking place during ontogeny, the adaptive character of developmental systems comes to the fore. The relevance of plasticity to the understanding of development and the origin of variations was rejected by the MS principally because of its Lamarckian connotations. As Lamarck's theory of acquired characters proposes, adaptions should be explained by the adaptive origin of traits in development. According to Waddington, acquired characters, once canalized, could become part of phylogenetic trends.

In this sense, what phenotypic plasticity also shows is the directedness of development, in this case, to an apt phenotypic result. Phenotypic plasticity, once it is understood not in genetic but in developmental terms, becomes another example of how developing systems are sensitive to their life conditions —their internal and external circumstances— and produce an adaptively-directed response. Both robustness and plasticity understood as interrelated processes (Bateson & Gluckman, 2011) are central in the contemporary view of teleological development.

To conclude this subsection, let us summarize the three basic insights in the thought of the four 20$^{\text{th}}$ century organicists discussed here.

1. As von Uexküll defended, goal-directedness resides in the *harmonic regulations of the U–I*. This should be the proper goal of the directedness of development, not adult states. Moreover, the U–I are relative to the organism, not independent from it.

2. As Russell explained in his characterization of goal-directedness, the creativity of development consists of the capacity for searching for alternative means towards a goal. The repertoire of variation is *limited* by the material constraints of development, while at the same time the developmental outcome is *not determined* by such constraints.

3. *A*s von Berlantaflly and Waddington noted, robustness and plasticity are two central developmental phenomena not reducible to the genetic level and evidencing the directedness of development toward adaptive phenotypic outcomes.

7.3 The contemporary view of teleological development

I opened this chapter by stating that while many proposals and scholars exist working on teleology at the physiological scale, the situation is not the same when it comes to development. To be sure, in my opinion there is only one solid and explicit account of teleological development that is aligned to and motivated by the DT. It has been proposed by Denis Walsh in a series of papers (Walsh, 2006a, 2006b, 2008, 2012a, 2012b, 2013a, 2014, 2018, 2019) and systematized in his book *Organisms, Agency, and Evolution*. As expected, Walsh draws on many insights from the works presented in the previous section, in addition to the view of agency introduced in the previous chapter. I will first present Walsh's view in Section 7.3.1. In Section 7.3.2 I shall offer a number of obseervations on the nature of teleological development that will help me supplement Walsh's view in Chapter 8. In Section 7.3.3 I will conclude by introducing Walsh's view on teleological invariance, and present a shortcoming of his analysis. This shortcoming opens the door to an alternative but complementary view of teleological development in Chapter 8.

7.3.1 The Agential Perspective of development

The notion of *Agential Perspective* of development was introduced by Sultan, Moczek, and Walsh (2022) to emphasize the fact that several biological phenomena would get much more adequate explanations if we treated development as a process guided by an agent. Even though Walsh does not use this label explicitly to refer to his view on teleology, I will adopt it because it emphasizes the importance of organismal agency and its relation with the biological perspective of the DT. Walsh's Agential Perspective turns around the following four notions which constitute "an interdefineable cluster" (Walsh, 2015, 211). As it can be noted, the ideas introduced in the previous, historically-oriented sections are implicit in Walsh's view:

Repertoire: A developing organism is capable to find different pathways toward an outcome within an adaptive repertoire of possibilities.

Affordances: Developmental responses are always the result of how an organism *experiences its environment*.

Agency: Organisms are agents capable of self-regulate their developmental trajectories.

Goals: Goals are the end state of the activity of goal-directed systems.

Repertoire

The notion of a *repertoire* is central to Walsh's view. As in Russell's proposal, a repertoire refers to the different outcomes that a system may achieve according to the constraints imposed by the system itself. An *adaptive repertoire* not only includes all the possible outcomes, but it also imposes an adaptive valuation over such outcomes, i.e. an adaptive repertoire establishes which outcomes are more apt than others. Walsh contends that organisms have a repertoire of possibilities and that choosing among them is the core of natural purposes.

The idea of a repertoire entails that not any variation is possible; the range of variation is constrained. The notion of constraint has always been fundamental in the DT, particularly in evo-devo (Amundson, 1994; Maynard Smith et al., 1985), and it was also present in Russell's view. We already saw in Section 5.1 that developmental resources make variation finite. Biological spandrels (Gould & Lewontin, 1979) and inherencies (Newman, 2021) show that the dynamics of development and the physical properties of developing systems operate as constraints of development. Therefore, *developmental constraints define the repertoire of developing systems*. What developmental constraints are exactly will be discussed in Chapter 8. In evo-devo, the idea of a repertoire is connected with that of a morphospace as it was introduced by Pere Alberch (1980; 1982; 1989). A morphospace represents the possible morphologies that a developmental system may achieve according to the constraints of development.

In some sense, the role of the adaptive repertoire is similar to the role of selection in externalist explanations of adaptive evolution endorsed by the MS. As noted in Chapter 2, the adaptive dimension of selection lies in the fact that external pressures pick up the fittest phenotype from a pool of different traits —the so-called design-space (Dennett, 1995): variation is required in order for selection to be possible. The case of the adaptive repertoire is similar. The adaptive dimension of development is possible insofar as developmental outcomes may vary under the same developmental constraints. The navigation of the repertoire towards an adaptive outcome is the landmark of goal-directed activity.

Like Waddington and von Bertalanffy, Walsh contends that "the plasticity and robustness of development are facets of an organism's purposivenes" (Walsh, 2021, 5). The plasticity and robustness of developmental outcomes are achieved by the stabilization and modification of developmental pathways toward the construction of an adaptive phenotypic outcome. "A system is persistent to the extent that

it can maintain, or resume, its trajectory by making compensatory changes when perturbed. In this respect, the phenotypic plasticity evident in organismal development is the very paradigm of goal-directed activity" (Walsh, 2006a, 441). The notion of adaptive repertoire is thus connected to those of robustness and plasticity: robust outcomes concern the production of the same outcome in different repertoires (i.e. under different developmental constraints), and plastic traits have to do with the production of different outcomes under the same developmental constraints.

This picture is clearly reminiscent of the famous epigenetic landscape introduced by Waddington (1957).[2] The epigenetic landscape was intended as a representation of the different pathways that a developmental system may transit towards a phenotypic outcome. The landscape is constituted by developmental constraints (i.e. material conditions of development) while a ball (the developing organism) is rolling down a hill. The trajectory of the ball is conditioned by the shape of the landscape. The organism rolling down the hill has many ways to achieve many outcomes, in a way that "biological form is biased by the capacity of organisms to originate and fix adaptive traits *because* such traits are conducive to survival" (Walsh, 2013b, 63; emphasis in the original), i.e. outcomes are adaptive responses.

Affordances

Walsh opts however for a different landscape metaphor: the *affordance landscape* (Walsh, 2012b, 2013a). The main idea is that what transits the epigenetic landscape should not be pictured as a ball rolling down, that is, as an object passively moved by external conditions. Here enters the scene the *experienced environment*. The epigenetic landscape is transited by an agent according to how the landscape is experienced by the agent itself, that is, not as a passive system moved by external forces but as an agentive system actively moved by internal forces. Walsh appeals to the notion of *affordance* to refer to how an organism experiences its environment in a way that determines different possibilities of action on it. While he takes the notion of affordance from Gibson's work in ecological psychology (Gibson, 1979), Walsh's affordances are more closely connected with von Uexküll's notion of *Umwelt*. Affordances are organismal-environmental constructs, such that the adaptive repertoire cannot be defined autonomously by the organisms but it is rather the result of the interactive nature of affordances. Thus, the agent navigates the landscape according to how it is perceived by the agent itself, and not as defined by external conditions, in order to choose an adaptive pathway towards an outcome within the adaptive repertoire.

[2] See Baedke (2013) for a nice historical analysis of the impact of the notion of epigenetic landscape.

Walsh's emphasis on affordances is an attempt to provide an alternative picture to the Explanatory Externalism of the MS adaptationism (Section 4.1.3). As explained in detail in Chapter 2, the autonomy of the environment defended by Explanatory Externalism leaves no role to developmental systems in the explanation of aptness. Aligned with Lewontin's insights concerning the centrality of organism-environment interactions, and the arrival of niche construction theory, affordances come to emphasize that the fit of organism and environment is not the result of an external force but of the way that the organism experiences an environment and adaptively responds to such experience. As Walsh puts it, "[a]daptive evolution is not most perspicuously described as the process of form solving adaptive problems set by the organism's physical environment, but as form creating and then responding to an ever-changing system of affordances" (Walsh, 2015, 178).

Agency

Walsh's view also pursues an agential view of teleology. The core idea of agency is that organisms are capable of regulating (i.e. adaptively modulate) their interaction with the environment in order to achieve a goal. Sultan et al. (2022, 1) define agency as "the capacity of living systems at various levels to participate in their own development, maintenance, and function by regulating their structures and activities in response to conditions they encounter". Clearly, this view opposes the gene-eye (or any replicator) view of development. In this context, the notion of agency stresses that the responsibility of developmental control lies on organisms, not on replicator units such as genes. Walsh's view is a natural child of the post-genomic view of development: organisms, not genes, self-regulate their developmental trajectories. Developmental agency also hinges on the idea that developmental outcomes are the result of multiple causal resources —genetic, genomic, intracellular, extracellular, and exogenous— integrated into the developmental system.

So agency is related to goals, repertoires, and affordances. Without a repertoire, there are no options to choose from different possibilities in relation to their adaptive value: we would not be able to see how "the agent negotiates its situation" (Walsh, 2015, 210). Without a goal, an agent's activity would not make sense: there would be no reason why the agent chooses one or another possibility. Without affordances, the activity of developing organisms would succumb to external forces.

Goals

Walsh's view of goals is relatively canonical: a goal is the end-state that a goal-directed process is directed toward. But in order to avoid circularity, he needs to explain what is a goal-directed system.

His explanation of goal-directed systems relies on a long tradition of empirical and conceptual research already presented in this thesis: cybernetics, systems biology, self-organization, adaptive systems, and autonomous systems, among others. This tradition allows him to conclude that goal-directedness "consists in the capacity of a system as a whole to enlist the causal capacities of its parts and direct them toward the attainment of a robustly stable end-point" (Walsh, 2015, 195). Goal-directedness is a property defined in terms of a whole-parts relationship in an organic system. The whole orchestrates its parts towards an adaptive outcome. Walsh's view, therefore, is an organicist one, where goal-directedness is assumed to be an emergent phenomenon, it is "an observable, gross behavioral property of a system" (Walsh, 2014, 206).

In sum, Walsh's Agential Perspective allows us to see development as a goal-directed phenomenon and teleological explanations of development can be accounted for with the conceptual apparatus presented here, "because there are agents, there are goals, means, norms, hypothetical necessity, and a special mode of explanation —teleology" (Walsh, 2018, 172).

7.3.2 Complementing Walsh's Agential Perspective

Here I would like to expound two brief remarks on teleological development. They should not be interpreted as challenges to Walsh, but rather as complementary ideas that will guide my analysis in the next chapter and which are required for a naturalized theory of teleological development.

The double way of being of organism

What is the goal of a development system? Are the goals of processes taking place at the developmental scale the same as the goals at the physiological scale? As a first approximation, since autonomous systems theory is about processes at the physiological scale, we can ask whether developing organisms are autonomous organisms. Or, following Nuño de la Rosa (2010, 294), we can pose the following question: "[c]an we really apply the organisational definition [of autonomy] to all the stages of a life-history?" To answer this, she analyzes the different developmental stages in animals (especially vertebrates). While we might say that at the zygote-stage we have an autonomous system (cells were taken to be the minimal case of autonomous systems), her analysis concludes that in the rise of

multi-cellular organization during further developmental stages many properties of autonomous systems are absent. For instance, in the cleavage and gastrulation stages, there are no clear boundaries between the developing system and its environment (a central requisite for individuality in autonomous systems theory). Moreover, a full-fledged and functionally integrated —i.e. operationally closed— system is not achieved up to the end of organogenesis, when we find different organs (constraints in an autonomous system) performing different functionalities. While this is an empirical issue that deserves further study, it does not strike me as too premature to conclude that it is an empirical mistake to directly transfer the properties of autonomous systems to developing systems.

Anyway, the question of whether developing organisms are autonomous organisms is not my central target here. Surely, this is an interesting issue, and some of the properties of autonomous systems would be discussed in Section 8.2 in the context of development. But I need not solve the issue in one direction or another in order to pursue my investigation. Maybe autonomy is present at some stages and not others, and this varies depending on the species. Maybe the properties of autonomous systems come in degrees and emerge at different ontogenetic stages. In Chapter 8, I will argue that teleological development is connected to the *agential capacities* of developing systems and I will present arguments in favor of the thesis that developing systems are agents, leaving open the question of the place of autonomy in development.

Having said so, my central aim here is a different one: to argue that *developmental goals are not about self-maintenance*. Autonomous systems theory posits that the goal of autonomous systems is that of self-maintenance. However, development seems not to be about maintaining a structure, but about changing structure; development is about change, not preservation; teleological physiology aims to maintain an organization in optimal conditions, but teleological development is about the creation of an organization. Organisms have a *double way of being* (Nuño de la Rosa, 2010, 295). These two ways of being, I will call *the changing way of being* and *the maintenance way of being*, respectively. The double way of being of organisms requires different kinds of goals. Developing systems transit different configurations during their lifespan. While the organism's structure remains relatively stable in physiological processes, this is not the case during development. A single configuration of matter in a specific developmental stage does not define the identity of a developmental system (as in the case of autonomous systems) to the extent that such a configuration will change during development. Then, what does? If different organizations of constraints arise during ontogenesis, then, as (Minelli, 2009, 67) said, "one could ask how many different 'lives' an animal may fit into its cycle".

The changing nature of ontogenetic processes becomes even more dramatic

when we examine carefully the complexity of certain life cycles. As Moran (1994, 574) proposes, complex life cycles are constituted by "discrete phases that exhibit contrasting morphological, physiological, behavioral, or ecological attributes", with probably the most drastic change in ontogenesis being that of metamorphosis. Moreover, Griesemer (2014, 2016) notes that scaffolding processes make the constitutive dimension of developing organisms even more complex. A resource may operate as a scaffold of development if it is part of the developmental system at a certain stage and it is central to achieving a particular outcome in another stage, but once the outcome is produced, the developmental resource may be pulled apart from the system —or in other cases hybridized with it (Balari & Lorenzo, 2018; Griesemer, 2014). Scaffolding adds complexity to the dynamics of development.

The double way of being of organisms is a consequence of different kinds of goals. One kind is about change, and the other is about maintenance. One is about modifying the system, and the other is about preserving the system. Certainly, organisms transit periods where changes are ubiquitous, and other periods where stability is predominant. However, it is relevant to note that both ways of being are not differentiated on time, but that there is an overlap between them.

First, it is not obvious that development ends at all. While stability is achieved at certain stages in life cycles, further modifications can arise in future stages. Even, as Minelli (2011) argued, aging could be taken as a developmental process where anatomical, physiological and behavioral changes still take place. Also, the open-ended character of development is also relative to the traits involved. While morphology may remain relatively stable once it is constructed, cognition appears to be the paradigmatic case of open-ended development (Balari & Lorenzo, 2014b). In sum, while in their life cycles organisms exhibit stages of constant change and other stages where stability is predominant, *developmental goals remain active even once a high degree of stability is reached.*

Concomitantly, *self-maintenance is also central to development*: developing embryos can die. Therefore, different parts of the developing system must fulfill their roles in order to maintain the system alive: while changes may be constant, the developing organism is also busy with the task of its self-maintenance. There is functional differentiation in traits that are crucial for the maintenance of the developing system, even if further changes occur in those parts. As Gilbert and Barresi (2010, 1; emphasis added) state, "one of the critical differences between you and a machine is that a machine is never required to function until after it is built. Every multi-cellular organism *has to function even as it builds itself*". The net result is that *both ways of being overlap.* Their boundaries are not determined by any single moment in the lifespan, but both ways are mixed, *with periods where change is predominant, and a periods where maintenance is primary.*

In sum, the issue of whether developmental systems are autonomous or not is

peripheral to my aims of naturalizing teleological development. The key notion is that of 'agency'. Notwithstanding, this brief inquiry into the place of autonomy in development has allowed us to underscore an important fact: development is not primarily about maintenance but about change. Organisms have a double way of being: one way is about self-maintaining the system at the edge of chaos, and the other is about changing and constructing their structure and functioning. While organisms transit periods of drastic changes and periods of stability, both ways of being overlap: organisms change while they maintain themselves alive.

Teleological development without adulthood

What is an organism? This is a difficult question which I will not address here (but see my observations at the end of Section 6.3.3). Following the insightful analysis of Nuño de la Rosa (2010), I wish to emphasize the fact that *developing organisms* remain underexplored in most studies about the nature of organisms. The vast majority of accounts are based on the notion of Darwinian individuals (Godfrey-Smith, 2009), understood as those organic entities that may be involved in selection processes, while other approaches typically focus on some notion of adulthood. It is therefore of paramount importance to distinguish between *adult-oriented* and *development-oriented* views of organisms. The former is the focus of most discussions concerning the nature of agency and teleology. As for the latter, however, it is clearly the case that "developing organisms are not explicitly discussed as exemplars of the concept of organism" Nuño de la Rosa (2010, 291). Moreover, as I argued in Section 7.2.1, developmental goals are not adult-oriented. Consequently, in our quest for a theory of teleological development aligned with the idea of organisms as the central unit of development, the notion of adulthood is not a central one.

Surely, we can recognize that in some periods of the life cycle the system exhibits more stability, and in other periods the system undergoes many changes. However, as explained, a theory of teleological development needs not (must not?) be based on the notion of an adult organism. Firstly, because goals are not about adult phenotypic states, and my proposal on developmental goals (to be presented in Section 8.2) makes no reference to adulthood. Secondly, the very notion of adulthood is very difficult to define, if it is definable at all. On the one hand, the stability of an organism (the cessation of change) is relative to the traits analyzed. Cognitive stability, reproductive stability and morphological stability, to name what may be taken as three of the milestones in the development of an organism, need not be achieved at the same time. On the other hand, the fact that a trait remains stable during a certain period is no guarantee that no changes will occur in the future. For our purposes, then, it would suffice to state that traits transit periods of stability and periods of change, without actually getting into an

argument about whether there really is an end of development or whether some notion of adulthood may be firmly characterized.

One of the problems of Walsh's proposal is that it is not explicit enough about the double way of being. Consequently, Walsh's Agential Perspective involving goals, affordances, and repertoires must be adjusted to account for developing organisms and their goals without endorsing any adult-oriented perspective. How should goals, repertoires, and affordances be understood in developing systems? I propose here a first approximation to this question:

Affordances: Affordances should be relative to the actual context of the developmental system. They cannot be about the environmental context of future adult phenotypes. The developing system cannot see the horizon of the epigenetic landscape (i.e. adult phenotypes). It is a road without a preconceived end. In this sense, we must adopt von Uexküll's insight and claim that development is not directed to a future end but to the current U–I.

Repertoire: The adaptive repertoire cannot be adult-oriented. The adaptive repertoire must be relative to each ontogenetic stage,[3] i.e. a repertoire of possible pathways that a developmental system might take at each step in ontogeny. Therefore, the adaptive repertoire cannot be understood as adult-oriented —i.e. as a repertoire of adult phenotypes. This does not mean that an adult-oriented repertoire is a misguided conceptualization. I just state that developing organisms cannot be directed to an outcome within an adult-oriented representation of possibilities.

Goals: I already argued (section 7.2.1) against any sort of preformationism, i.e. against the idea that developmental ends (goals) are adult-oriented. End-states in development are not about 'the end of development'. Final causes are not about the termination of development. Rather, the end-states of development, and affordances and repertoires too, should be relative to each developmental stage. Development is an open-ended process.

7.3.3 Naturalization

In Section 7.3.1 I introduced Walsh's Agential Perspective which I complemented with some remarks that need to be taken into consideration in our understanding of teleological development (Section 7.3.2). One more important thing needs to be done, however: to evaluate the scientific validity of teleological explanations in

[3]The notion of a developmental stage will be clarified later on. For the present purposes it suffices to see development as a chain of changes through time such that at each point in time we have an ontogenetic stage.

development under Walsh's proposal, or, in other words, we still need to assess the extent to which teleological explanations in science are legitimate or not. This is my task in this section, where I first present Walsh's strategy for naturalization: purposive invariance. Next I will pinpoint a shortcoming in Walsh's analysis. This shortcoming, together with the comments exposed in the previous subsection, will guide my proposal in the next chapter, in addition to work as an aid to justify the need for complementing Walsh's Agential Perspective on teleological development.

Explanatory validity: Purposive invariance

Walsh's view rests on an novel view of development beyond the 20$^{\text{th}}$ century legacy and attempts to explain how teleological explanations are both explanatory necessary and legitimate in biology. As he states, "[r]ehabilitating teleology in evolutionary biology requires demonstrating three things: (i) that there are purposes in nature, (ii) that purposes make a difference to evolution, and (iii) that purposes can figure in genuine scientific explanations" (Walsh, 2021, 3). The first point is ontological. While my concerns here are mainly epistemological, I cannot deny that hints of teleology can be recognized in the agentivity, the adaptivity, and the plasticity of living beings, which seem to be legitimate elements of the structure of the world. The second point will be discussed later on in Chapter 8, once my proposal is sketched in connection with Walsh's one. It should not come as a surprise that a close connection between teleological development and adaptive evolution is necessary to understand how the causes of evolution (ontogenetic causes) can be made present in teleological explanations of development, i.e. how adaptive evolution results through adaptive ontogenetic processes explained in teleological terms. But, what about the last issue? It appears that teleological explanations are genuine scientific explanations of development insofar as Walsh's proposal is aligned with contemporary views defended by many biologists. However, we did not answer important questions concerning the validity of teleological explanations in development.

To account for the third issue —that purposes are part of genuine scientific explanations— Walsh provides a meta-analysis of scientific explanations. He first states which are the minimal necessary and sufficient conditions for treating a scientific explanation as legitimate, and then he argues that teleological explanations meet such conditions. According to him, there are two requisites that scientific explanations must fulfill: *invariance relations* and *elucidating descriptions*. Let me explain them in turn to see whether teleological explanations actually comply with them.

Concerning the first requisite, Walsh systematically appeals to the same strategy (Walsh, 2006b, 2012a, 2014): to provide an analysis in terms of invariance explanation based on the model of causal explanations provided by Woodward.

Thus, "by appeal to the apparatus of invariance explanation" he shows "that biological purposiveness provides genuine, ineliminable biological explanations" (Walsh, 2006b, 771).

Woodward's 'interventionist' strategy (Woodward, 2002, 2003) is based on a counterfactual analysis of the different variables involved in a scientific explanation. According to his model of invariance, in an explanation that posits a correspondence between X and Y to see whether X causes Y, we should evaluate the counterfactual situation where X does not take place or is modified, and check whether Y varies or not, or, in other words, we need to check the variance or invariance of Y under intervention on X. Therefore, "as a rough approximation, a necessary and sufficient condition for X to cause Y or to figure in a causal explanation of Y is that the value of X would change under some intervention on X in some background circumstances" (Woodward, 2003, 15). Walsh takes Woodward's model as a successful way to understand evaluative relations in scientific explanations, and then he applies it to his view of teleological development:

> Remarkably, the same kind of [invariance] relation holds between a goal and the means to its attainment. A goal is a counterfactually robust kind of difference-maker. For any system with goal e_1 that produces an event m_1, that is conducive to e_1 under actual conditions, under different conditions it would produce a different event m_2 such that m_2 would, in those conditions, conduce to e_1. Under similar circumstances, were the system to have had a different goal, e_2, then it would have produced a different event, m_3, conducive to e_2. (Walsh, 2014, 207)

That is, the same means would not occur in the same way if goals are intervened. The ontogenetic trajectory (means) to achieve outcomes varies under different developmental goals. If the developmental trajectory Y directed towards the phenotypic outcome X could be explained teleologically, and *teleological explanations are causal explanations*, then intervention on X (developmental goals) would lead to variations on Y (developmental means). Shortly, changing goals changes means.

The second requisite appeals to an elucidating description. Elucidating descriptions are part of any causal/mechanicist explanation: a causal/mechanicist explanation is not just about the fact that two phenomena are causally related (exhibit invariance relations), but we must describe the way in which the phenomena are related: we must elucidate their causal connection. Walsh notes that mechanistic explanations describe how the causes of certain processes or activities *produce* an effect. And these kinds of explanations are not merely descriptive but also elucidate how the relation between cause and effect takes place. As it is usually assumed, the paradigm of mechanistic explanations is a bottom-up approach

to a certain phenomenon, where an analysis is carried out of how different parts of a system interact to *produce* a particular effect.

Do teleological explanations provide elucidating descriptions? According to Walsh, they do. However, there is a crucial difference with mechanistic explanations. Mechanistic explanations elucidate how a cause produces an effect, but "when we offer a teleological explanation, we describe the way that the mechanism or cause in question *conduces* to the goal" (Walsh, 2012a, 178; emphasis in the original). And he adds: "Conducing, like producing, is a causal relation; means conduce to the attainment of goals only when means cause the attainment of goals". In this view goals *require* means to be fulfilled. Therefore, the elucidating character of teleological explanations is to explain how goals are fulfilled as a consequence of how means conduce to goals. In this sense, like mechanicist explanations that have their own vocabulary —e.g. binding, bending, pushing, attracting, opening (Walsh, 2015, 197-198), teleological explanations also have their own expressions, such as 'in order to', 'for the purpose of', or 'for the sake of'. The difference is that the mechanicist vocabulary invokes production relationships (cause-effect), while teleological explanations concern conducting relationships (means-ends).

Hence, purposive invariance is different from 'mechanistic' invariance: teleological explanations relate means and goals while mechanistic explanations relate causes and effects. Teleological elucidating descriptions explain how goals *conduce* to means, while mechanistic elucidating descriptions explain how causes *produce* an effect. Despite the differences, the relevant take-home message is that teleological explanations comply with the two central ingredients of causal explanations: "teleological explanations have the same form as causal-mechanistic explanations. They comprise two features: (i) a counterfactually robust invariance relation and (ii) an elucidating description" (Walsh, 2012a, 179).

Before concluding this section, we need to make explicit a further point. Walsh naturalizes teleological explanations by showing that they possess similar explanatory properties as mechanistic explanations. At the core of this solution lies the idea that not all scientific explanations are causal. In his view, teleological explanations would be an example of this. What his analysis shows is that we can explain means in terms of goals without presupposing that goals cause means (this would involve backward causation). The explanatory relations between goals and means exhibit the same meta-theoretical properties as the explanatory relations between causes and effects, so teleological explanations should be considered as legitimate as causal/mechanist explanations are. In my opinion here lies the main weakness of Walsh's analysis.

On having a goal: a difference maker

I see Walsh's justification of the validity of teleological explanation —i.e. purposive invariance— on the right track but not fully adequate. Recall that his conclusion is that teleological explanations are not causal but still scientifically legitimate: by appealing to the apparatus of invariance relationships, he shows that teleological explanations exhibit the same explanatory structure as causal/mechanicist explanations.

I think that this strategy is adequate, but that Walsh's analysis is not accurate enough. The first hint that this is so comes from the fact that he concludes that teleological explanations are not causal, and in doing so, he appeals to the tools of invariance relations, which are tailored to assess the legitimacy of *causal* explanations, not scientific explanations in general: Woodward's model is a model about causal relations. Besides, I think that there is a much deeper flaw. The main reason that leads Walsh to argue that teleological explanations are not causal explanations is the confusion between a *goal* (an end-state) and *having a goal* (an actual state of the organism that precedes both means and goals). I will argue that purposive invariance is not a relation between goals and means, but a relation between having goals and means.

There is a straightforward way to argue that there are no counterfactual dependencies between goals and means: counterfactual dependencies apply to relations between two phenomena, X and Y, where X temporally precedes Y. X cannot be the goal (the end-state) because the goal does not come before Y (the means). Something in the present cannot make a difference in past events: this is the problem of backward causation in teleological explanations. Walsh is well aware that backward causation is out of play and he is at pains to show that in teleological explanation there are invariant relationships and difference makers. But which ones?

I believe that real invariance relations in teleological explanations are not between goals and means, but between having a goal and a means. Having a goal is not the end of the goal-directed activity but the beginning. We can clearly see that a counterfactual dependence exists between means and having a goal. If system A had different goals, then the means would be distinct. Having a goal precedes means, and means precede goals. I think that this is indeed what Walsh is trying to show in his analysis of purposive invariance: "*a goal* explains its means because were the system in question *to have the goal* then the means would obtain, and were it *not to have the goal* the means would not have occurred. In this way, goals explain" Walsh (2014, 207; emphasis added). This captures the key idea that means take a particular form because of the goals the system is pursuing; i.e. that developmental pathways are a consequence of developmental goals. Having a goal is a causal notion that affects means. An organism with a different goal

(having a goal) would do different things (means) to achieve the goal (end-state). In contradistinction with Walsh's formulation that states that goals are difference makers, I am suggesting that having a goal is the real difference maker. Having a goal is a difference maker that makes a difference and, usually, it makes a good (adaptive) difference.

As I pointed out above, I believe that the distinction between having a goal and goals is already present in Walsh's analysis, but it unfortunately did not make its way to his conclusion. In other words, Walsh's analysis of counterfactual dependencies *does take* having a goal as the variable, not the goal itself, but, and here is the problem, Walsh concludes that invariance relations are between means and goals:

> For any system with goal e_1 that produces an event m_1, that is conducive to e_1 under actual conditions, under different conditions it would produce a different event m_2 such that m_2 would, in those conditions, conduce to e_1. Under similar circumstances, were the system to have had *a different* goal, e_2, then it would have produced a different event, m_3, conducive to e_2. The relation between an end (goal) and its means is thus an invariance relation. (Walsh, 2015, 198; emphasis in the original)

The confusion between goals and having goals appears in two places in this quote. First, Walsh states that a system with a goal (e_1) *produces* a particular means (m_1). He also explicitly states that the verb *produce* is the one that appears in causal relations between causes and effects, not between goals (effects) and means (causes). He is using a causal vocabulary to describe the relationship between goals and means. I believe that what Walsh is actually showing is that there exist causal relations between having goals (causes) and means (effects).

The second source of confusion concerns the fact that his analysis involves both having a goal and the goal (end-state) itself: "any *system with goal* e_1 that produces an event m_1, that is conducive to e_1"; or also: "were the *system to have had* a different goal, e_2, then it would have produced a different event, m_3, conducive to e_2". What is crucial here is that a system with a goal is not the same as the goal (end-state). I think that this difference is explicit in these quotes: having a goal produces means, and means conduce to goals. Walsh is indeed showing that there are counterfactual relationships between means and having a goal. I think that this is the correct picture, but Walsh's muddling of having goals with goals leads him to the mistaken conclusion at the end of the quoted paragraph: "The relation between an end (goal) and its means is thus an invariance relation". This is not correct. The relation between ends and means is not invariant; rather it is the relation between having goals and means that is invariant. If we get rid of

this confusion, I believe that Walsh's analysis actually sheds light on the *causal structure of teleological explanations*, which I graphically outline in Figure 7.1.

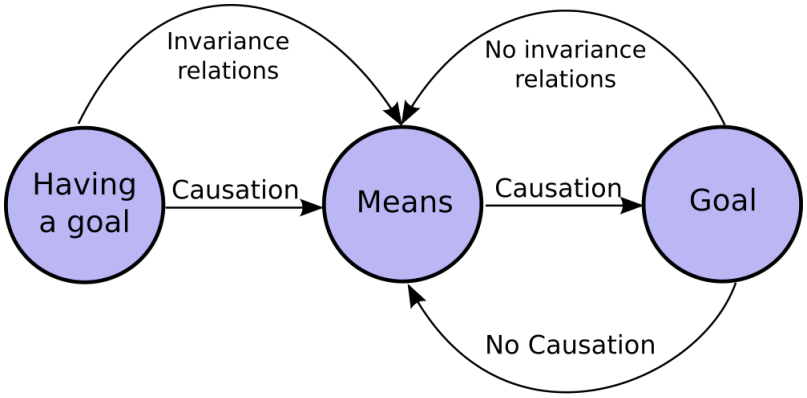

Figure 7.1: The causal structure of teleological explanations.

The relation between having a goal and a means is causal, but this is not necessarily to be read in mechanistic terms. The difference between mechanistic and teleological explanations is not that the latter invert the order of cause—effect relations. Indeed, this is what, as I see it, the analysis of Walsh shows once we remove the confusion between goals and having goals. The actual difference is that the way in which having a goal produces a means is not the same way in which a cause produces an effect in mechanistic explanations. In other words, the relation between having a goal and a means is not mechanistic: it cannot be understood in a bottom-up fashion through a parts-whole analysis. Now, a new question arises: *how does having a developmental goal produce a means to conduce development to the attainment of the said goal?* Or, to put it differently: how should we understand the different items in the causal structure of a teleological explanation?

Towards Agential Teleosemantics

In order to answer the question posed at the end of the preceding section, I will introduce *Agential Teleosemantics* in Chapter 8. There, I will argue *that teleological explanations of development are a sort of intentional explanation.* Obviously, this idea requires some qualifications. I am certainly not going to argue that zygotes have mental representations. Rather, I will argue that *the causal structure of teleological explanations is identical to the causal structure of intentional explanations.* The idea, therefore, is that while goals do not have causal powers, intentional states do: they can produce means that are directed to a goal.

The strategy of positing some sort of intentionality to account for the temporal dimension of natural teleology is not entirely new and it is already present in other accounts. For example, Neander (2018, 65) analyzes the structure of teleological explanations taking a perspective that is similar to mine: teleological explanations (TE) appear to have the structure shown in TE1, but on a different analysis —like the one pictured in Figure 7.1 on the preceding page— they eventually acquire the form shown in TE2:

TE1: A goal at t_3 explains a means at t_2.

TE2: A representation at t_1 of a goal to be achieved at t_3 explains a means at t_2.

It is therefore important to distinguish between Type 1 Teleological Explanations (TE1) —goals explain means, as proposed by Walsh, from Type 2 Teleological Explanations (TE2) —having a goal produces means to conduce the system to a goal.

Most proposals concerning a teleological view of nature take the form of TE2. The teleonomic strategy of the MS, for example, also falls under TE2. Paley's natural theology or Plato's demiurge also posit representations of the goals (in an intelligent designer) before goals occur. Significantly, the kind of explanation involved in TE2 is causal. Therefore, having a different goal (variation at t_1) *causes* a different mean (variation at t_2). This is the reason why teleological explanations exhibit invariance relationships. TE2 is the form of teleological explanation I shall appeal to in the next chapter when presenting Agential Teleosemantics.

A shortcut without eliminativism: von Baer revisited

Let me conclude by clarifying a point that may be helpful to understand my position. The shortcut strategy for teleological explanations is usually eliminativism: teleological explanations do not provide genuine scientific explanations, they are just a shortcut for genuine scientific explanations (cf. Section 6.1). Teleological language can be reduced to non-teleological language in such a way that teleological explanations are not valid but only useful in scientific practice. I defend the shortcut strategy but without endorsing eliminativism. My claim is that *TE1 is a shortcut to TE2*. In other words, that TE1 can be reduced to TE2. Clearly, my strategy is not eliminativist because TE2 cannot be reduced to a non-teleological explanation.

I think that the distinction between TE1 and TE2 helps appreciating what falls in and what falls out of natural teleology; that is, which teleological expressions are naturalized (i.e. can be integrated into genuine scientific explanations) and which ones deserve just a heuristic treatment (i.e. are useful tools for scientific practice).

This task is a mandatory one insofar as teleological language has traditionally been charged with negative connotations in science. So it is important to pinpoint what will *not* be naturalized.

Fortunately, the hard terminological disputes were already settled some time ago. As von Baer argued —and von Uexküll and Russell followed suit— teleological explanations in biology involve goals (*Ziele*) but not purposes (*Zwecke*). My point is that von Baer's proposal can be incorporated to my analysis of purposive invariance and to the difference between TE1 and TE2.

Walsh's analysis of purposive invariance grants him the possibility to explain how goals explain means. In doing so, he argues that teleological explanations elucidate the relationships between means and goals by showing how means conduce to goals. Conductive relationships, Walsh claims, enable the use of teleological "locutions like 'in order to', 'for the purpose of', 'for the sake of' … to signify that the effect in question is a goal" (Walsh, 2015, 198). However, if TE1 is reduced to TE2, all the explanatory work is carried out by how having a goal orchestrates means towards a goal. The strategy of Walsh for naturalizing some of the aforementioned locutions fails to the extent that it rests on a flawed analysis of invariance relations in teleological explanations, as previously argued. It is certainly not so easy to trace a clear and distinct line between what can and what can not be naturalized. My point here is that —motivated by von Baer and by Figure 7.1 on page 232— the central explanatory notion that involves a teleological locution is that of a system *having a goal*, and that, as von Baer believed, other teleological expressions may be interpreted just an 'as-if' talk, as it is the case with *Zweck* (purpose). Under this interpetations, speaking about 'Zwecke' is simply a different way to talk about 'Ziele'. In the next chapter I shall argue that having a goal can be naturalized and, more importantly, that it underlies the central teleological phenomena of development.

7.4 Summary

As I advanced at the beginning, the aim of this chapter was not to offer a comprehensive view of teleological development. Here, my goal was simply to set down a number of important observations regarding the developmental time scale and teleological explanations (Section 7.1), in order to highlight several key ideas already advanced during the 19th and 20th centuries (Section 7.2) and to present the contemporary view of teleological development within the Agential Perspective defended by Denis Walsh (Section 7.3). As in the case of the physiological scale, the naturalization of teleology is in this case connected to a number scientific improvements in the understanding of developmental processes. In this sense, as many other have noted before (e.g. Gilbert & Sarkar, 2000; Oyama,

2001, 2010), different concepts —such as teleology, holism, emergence, agency or even epigenesis— have had fairly negative connotations in biology that deserve to be cleared out.

In Section 7.2.1 I analyzed the preformation–epigenesis debate, considering both its old (morphogenetic) and its new (informational) versions. I derived three results from this analysis:

(i) That in biology questions about teleology can be translated to questions about information.

(ii) That a developmental theory of inheritance and the notion of ontogeny of information are important tools to fight against modern preformationism.

(iii) That goals are not preformed —i.e. that developmental goals are not adult-oriented.

Next, in Section 7.2.2, I introduced the teleomechanistic program in order to present von Baer's central insights: firstly, his terminological distinction between purposes and goals, and his defense of the latter against the anthropomorphic and non-naturalist connotations of the former; secondly, his claim that goal-directedness underlies the harmonic relation existing among the parts of the organism and between the organism and its environment throughout ontogenesis; and finally, his appeal to a vivid musical metaphor to understand how such regulations could be possible.

I finished this section by underlining a number of important insights from organicism, starting with von Uexküll definition the U–I relative to the organism's experiences and his suggestion that the harmonic regulations stressed by von Baer should be understood as directed toward the U–I and not to adult states. Thereafter, I turned to Russell and his emphasis on the creativity of development as a mark of goal-directedness, understood as the capacity of following different routes towards the same end (different means for the same goal), in such a way that the material basis of development constraints but not determines the possible repertoire of outcomes. Finally, I briefly examined von Bertalanffy's definition of equifinality as a characteristic of open systems in order to stress the directedness of development towards the same outcome, and Waddington's theory which singles out plasticity as a cardinal element in the adaptiveness of development.

All these ideas were reappraised in Section 7.3.1 and contrasted with a contemporary view of teleological development: Walsh's Agential Perspective. Recall that Walsh's view hinges on four interconnected conceptualizations: repertoire, affordances, agency, and goals. In Section 7.3.2 I set forth a number of observations on teleological development that will guide my proposal in the next chapter.

These remarks are motivated by the double way of being of organisms: not everything is about maintenance. Organisms also have a *changing way of being*, and developmental goals are about the adaptiveness of such changes. Recognizing that organisms also have a changing way of being required to re-frame the conceptual apparatus of Walsh for developing organisms without any reference to adulthood.

I concluded in Section 7.3.3 by analyzing Walsh's project of naturalization, where the legitimization of teleological explanations is accounted for by providing a meta-theoretical analysis to establish what are the necessary and sufficient conditions to assess scientific explanations, and argued that teleological explanations actually comply with them. Walsh identifies two conditions that are met by teleological explanations. On the one hand, teleological explanations show invariant relations: if goals change, means also change, such that means are made counterfactually dependent on goals. On the other hand, teleological explanations elucidate how means conduce towards the achievement of a goal. Teleological explanations, Walsh concludes, exhibit a specific kind of invariance relations —a.k.a. purposive invariance— that make teleological explanations legitimate in science.

I argued that the naturalization project of Walsh is on the right track but inaccurate. Particularly, his argumentation conflates the notions of having a goal with the goal itself. I observed that this distinction is nonetheless explicit in Walsh's analysis, but not in his conclusion. Invariance relations exist not between goals and means, but between a system having a goal and the means that the system elicits as a consequence of having goals. Accordingly, means are counterfactually dependent on a system having a goal, not on the goal itself. In this sense, the structure of teleological explanations is not about goals explaining means, it rather is about having goals that cause means that in turn cause goals: this is *the causal structure of teleological explanations*.

Agential Teleosemantics, to be introduced in the next chapter, is directly motivated by the causal structure of teleological explanations. I will argue that teleological explanations in biology are intentional explanations: both teleological and intentional explanations have the same causal structure. So the aim of the next chapter is to analyze the causal structure of teleological explanations in intentional terms. From now on our keywords shall be: information and agency.

Chapter 8

Agential Teleosemantics and biology

In this chapter, I propose my own view on teleological development: Agential Teleosemantics. It is complementary to Walsh's Agential Perspective, and it is also motivated by the precursory work of a number of figures in the history of theoretical biology, while also being influenced by many insights from other contemporary proposals.

The chapter is organized as follows. The presentation of Agential Teleosemantics is split up into two sections. First, in Section 8.1 I shall introduce the technical notion of biological information advanced by Paul Griffiths and Karola Stotz, and I will connect it with the eco-devo view of development and with biosemiotics. Next, in Section 8.2 I will link biological information with developing agents. This task requires defining all the conceptual apparatus of teleological explanations, such as developmental constraints, repertoires, and the experienced environment, in terms of biological information. To conclude, in Section 8.3, I will make further clarifications concerning different issues around teleological development and address different points related to the explanatory adequacy of Agential Teleosemantics.

8.1 Information in development

The aim of this section is to introduce an informational account of development. As it was made explicit in the discussion of the preformation-epigenesis debate, the notion of information is fundamental for my proposal on teleological development. The conjunction of von Baer's claim that "natural forces which are not directed to an end cannot produce order" (von Baer, 1886b, 88) with Oyama's idea of information as "the cause that imparts order and form to matter" (Oyama, 2000b, 3), information becomes the directive force in living systems responsible

for organizing matter.

The notion of information, moreover, encourages a direct connection with intentionality. Since the cognitive revolution, cognition has been understood in terms of processing of representations and processing entails (at least) two things: information and normativity. The technical notion of biological information to be introduced in this section will be instrumental for my defense of a specific sort of intentional explanations in development or, more specifically, the idea that the elements that constitute the causal structure of teleological explanations can be defined in intentional terms.

I will appeal to three different research areas that will help building an informational view of development: the post-genomic view on biological information developed by Paul Griffiths and Karola Stotz, eco-devo, and biosemiotics. I will end by connecting these views with Walsh's analysis of teleology, principally by focusing on the intersection between these different proposals, which is the analysis of causal explanations based on Woodward's framework modeled with graph theory.

8.1.1 Biological information

This subsection introduces the informational theory of development expounded by Paul Griffiths and Karola Stotz. Their account will be linked with different important elements of my proposal. First, it will prove useful in the analysis of signaling systems in informational terms along the lines of some recent proposals within eco-devo; second, in pursuit of one of biosemiotics' main goals, it will help specify the notion of interpretation of signs; and third, it will contribute to grounding the causal bases of teleological explanations, as Walsh intends. All these elements will be introduced in this section. In addition to these connections, in Section 8.2 I will appeal to the notion of biological information to define the different elements participating in teleological explanations of development. To round it all off, in Chapter 9 I will explore another connection enabled by this informational account: the role of information as a potentially unifying factor of intentional explanations in cognitive science and teleological explanations of development.

Plenty of proposals make some use or another of the notion of information in connection to development. Francis Crick's Central Dogma is a celebrated example of this. Not surprisingly, here I will notwithstanding appeal to an approach that fits better with the aims and tenets of the DT. Paul Griffiths and Karola Stotz, two important thinkers for the setting up of philosophical underpinnings of the DT, have been pursuing an informational view of development in the last years and it is in this work that I will mainly focus on (Calcott, Pocheville, & Griffiths, 2020; Griffiths, 2013, 2016, 2017; Griffiths et al., 2015; Griffiths & Stotz, 2013; Stotz, 2019; Stotz & Griffiths, 2017). In what follows, I will use 'G–S' to refer generically

to the work of Griffiths, Stotz, and other colleagues on the role of information in development.[1] While initially the attitude towards information in biology within Developmental Systems Theory was one of suspicion inasmuch as it was perceived as connected to a preformationism of sorts (e.g. Oyama, 2000b), G–S's proposal aims at turning information into a valid epistemological tool for understanding development beyond the vestiges of preformationism.

Crick Information

G–S appeal to Crick's notion of information —a.k.a. 'Crick Information'— in order to determine to what an extent other non-genetic resources also possess this kind of information. They present their proposal as a bottom-up strategy, defining the kind of information present in genes (specifically, in coding regions), and then extending this view to epigenetic and exogenous developmental resources. G–S define 'Crick Information' as the notion of information involved in Crick's formulation of the Central Dogma: "Crick understood information twofold, as (1) precise determination, and (2) the transfer of this biological specificity from one molecule to another" (Stotz, 2019, 325).

In Crick's Central Dogma (also known as the 'Sequence Hypothesis'), the precise determination of biological information relates to how DNA sequences determine the protein outcome. This determination lies at the core of the Dogma. Biological information about protein outcomes resides in DNA to the extent that which specific protein products will be produced is determined by amino acid sequences. The unidirectional flow of information in development, as the Dogma states, implies a direct mapping between DNA sequences and protein products. Concerning the second point, as Stotz and Griffiths (2017, 366) point out " 'information' is a way to talk about specificity". Developmental causes that provide biological specificity are distinct from efficient causes. Specificity is analyzed in terms of the *information* provided by a developmental cause, while efficiency is analyzed in terms of *matter and energy*. As explained by Stotz (2019, 325), "distinguishing 'matter and energy' from 'information' corresponds to the distinction between the efficiency and specificity of a molecular process". In the Central Dogma, developmental specificity refers to the capacity of DNA of specifying developmental outcomes by the unfolding of the evolved information coded in amino acid sequences. While many causal resources participate in development, DNA is the unique source of specificity (information) while other, epigenetic causes, are only

[1] Griffiths and Stotz (2013, 144-145) distinguish two kinds of informational talk in genetics. One is *information in genes* and the other is *information about genes*. Both are certainly important but our emphasis will be on the former, that is the information about development (if any) that genes possess. The latter regards different issues in bioinformatics, bioethics, or medicine concerning the use in each discipline of information about genes.

sources of efficiency (matter and energy). In this sense, the distinction between specificity and efficiency maps into the distinction between genetics and epigenetics. A developmental specifier provides information about which developmental outcome must be constructed. Accordingly, biological information is connected to how DNA sequences precisely determine outcomes by carrying specific information about which outcomes must be produced: "the essential idea to take home is that information in biology is a way to talk about a specific kind of causal relationship, that of specificity: DNA molecules carry information in virtue of their ability to specify precisely the linear structure of the resultant proteins" Stotz (2019, 325).

As expected, the main problematic point of Crick's view is not his notion of information but his claim that DNA determines outcomes, the idea that no further information is needed to produce the proper developmental outcomes: "When Crick advanced the sequence hypothesis and Central Dogma he assumed that the sequence of the gene not only *precisely* determined the sequence of the product but also *completely* determined it" (Griffiths, 2016, 83; emphasis in the original). Such determination in the post-genomic era is typically assumed to be *distributed* (Stotz, 2006a) among different causal factors and, thereby, Crick's information extends beyond DNA coding regions: "Crick's approach to information has natural extensions to non-coding regions of DNA, to epigenetic marks, and to the genetic or environmental upstream causes of those epigenetic marks. Epigenetic information cannot be reduced to genetic information" (Griffiths, 2017, 1). As a consequence of this, the distinction between efficiency and specificity does *not* directly map onto a distinction between genetic and epigenetic causes of development. The many sources of non-genetic specificity demand "additional specificity of a kind not captured by the original 'sequence hypothesis' " (Griffiths, 2016, 83).

In Crick's view, specificity is exhausted in coding regions of the DNA. But the notion of 'Crick information' that G–S appeal to is not committed to any sort of genetic determinism. Rather, it is just based on the idea of one developmental resource providing information that up to a certain degree and in collaboration without other resources, specifies which outcomes may be constructed; see below for a detailed exposition of the G–S view. Crick Information is, therefore, related to developmental control. A high degree of developmental specificity exerts a *fine-grained control* of development: changes in developmental specifiers entail changes in developmental outcomes. The central point in post-genomics turns around the distribution of such developmental specifiers as a direct answer to the question about which entity controls development. In Crick's view, the precise determination of DNA sequences "expressed how fine-grained changes in the DNA/RNA sequence *cause* fine-grained changes in protein structure" Stotz (2019, 324; emphasis in the original). If we move beyond the molecular gene, we reach the thesis of distributed specificity —and also of distributed developmental control— aligned

with the view promoted in the post-genomic era.

Be that as it may and despite the differences, the common ground is about how biological information should be understood. The core of Crick's proposal that G–S adopt is that biological information concerns the specificity of developmental outcomes as a consequence of the fine-grained control of different developmental resources. With this preliminary idea in mind, I will present in detail their proposal, which, while based on the notion of Crick Information (developmental specificity), offers a much more sophisticated picture than Crick's Central Dogma, partly because the democratic view of developmental specificity that it advances is much richer than the genetic reductionist one.

Information and Causation

What makes G–S's proposal attractive is not merely that it fits with the view of the DT about the nature of development but also that it is connected with an analysis of causation.[2] The core of the G–S view is their definition of biological information in terms of *causal specificity* by appealing to Woodward's framework on causal explanations. So if information is taken to be the cardinal element in teleological explanations of development, and information is defined in causal terms, an informational account of teleology may be appropriate to dealing with the causal basis of teleological explanations. G–S "propose that causal relationships in biological systems can be regarded as informational when they are highly causally specific" Stotz and Griffiths (2017, 374), insofar as "causal specificity is closely related to the idea of biological specificity" Stotz and Griffiths (2017, 371).

We already met the main tenet of Woodward's interventionist theory. One variable may be said to cause a certain effect if manipulation on such a variable brings about a different effect. So in any causal explanation in science, we can in principle know whether the elements involved in the explanation stand in a causal relationship or not by verifying when intervention entails variation in the outcomes. Within this framework it is therefore possible to define causal specificity in connection with fine-grained control, and thereafter and in connection to this, the notion of (biological) informational specificity.

An important step to achieve this goal, is to realize that developmental phenomena are not the product of a single cause: phenotypic outcomes are not a matter of switching on or off some developmental resource; the relation between developmental resources and developmental outcomes is not one-to-one but a one-

[2]The connection between biological information and causality has also been acknowledged by Woodward (in, e.g. Woodward, 2010) and by Johansson (2007, 84) who states that "we can't get any information from a system without interacting causally with it […] information is a causal process"; see also Moreno and Ruiz-Mirazo (2011), Shea (2007b), Jablonka (2002) and Pharoah (2020) for discussion.

to-many mapping. There is a gradient of different causal influences that a developmental resource has in the dynamics of development: "causes differ in the degree to which they are 'specific' to their effects" (Griffiths, 2016, 81). In this sense, if a developmental resource exercises a fine-grained control on a developmental outcome, such a resource is a source of causal specificity. As Griffiths puts it, "in informal terms the cause can make the difference between *many different states* of the effect and can be used to exercise fine-grained control over that effect" (Griffiths, 2016, 82; emphasis added). As expected, we may then be able to determine whether a developmental resource is a causal specifier by intervention, because the degree of specificity that a developmental resource might have is relative to the degree of change in the effects the resource produces under intervention: "specificity refers to the fine-grained control that an intervention might have, controlling a gradient of change, rather than a simple on-off switch" Stotz and Griffiths (2017, 372).

Just as information is distributed among different resources —as the defenders of the Parity Thesis support— specificity is also distributed. Applying the idea of distributed causal specificity to development emphasizes the fact that developmental processes result from different developmental resources. The metaphor of a tuning dial of a radio, due to Calcott et al. (2020), illustrates that the degree and distribution of causal influences can be extended to development: there is no developmental cause that works as the on/off switch of a radio, but rather the different resources operate as tuning dials of development depending on their causal specificity:

> The intuitive idea behind causal specificity can be illustrated by contrasting the tuning dial and the on/off switch of a radio. Both the tuning dial and on/off switch are causes (in the interventionist sense) of what we are currently listening to. But the tuning dial is a more specific cause, as it allows a range of different music, news, and sports channels to be accessed, whilst flipping the on/off switch simply controls whether we hear something or nothing. (Calcott et al., 2020, 246)

Measuring causal specificity in informational terms

So far, I merely sketched the idea that biological information is related to specificity and that, accordingly, specificity in developmental systems may defined in causal terms. The net result of this move is that causal specificity can be understood in informational terms, and therefore, we should in principle be able to measure the causal contribution of a developmental resource in terms of the information it conveys to development. This idea is just an extension of the previous analysis of causation: "causal relationships in biological systems can be regarded

as informational when they are highly causally specific" (Stotz & Griffiths, 2017, 374), or alternatively, "*biological information is a substantive* [i.e. highly specific] *causal factor in living systems*" (Stotz, 2019, 323; emphasis in the original). The main contribution of G–S is to define a measure of how much specificity a causal variable provides by appealing to informational theory; see Griffiths et al. (2015) for a detailed presentation.

The notion of information at play is the conventional one introduced by Shannon and Weaver (1949). I will not get into mathematical details because the theoretical aims of my analysis are largely independent of their precise formalization. It is sufficient for us to keep in mind that (Shannon) information has to do with variation on the levels of uncertainty. Information decreases uncertainty by changing the probability of the possible states of the world. For instance, if I do not give you any information about the date of my birthday, any day of the year is equally probable (1 out of 365); if I inform you that my birthday is in May, uncertainty decreases, now the probability is 1 out of 31; if I say that my birthday is in the first week of May, now the probability is 1 out of 7; if I eventually say that my birth is on the 7^{th} of May, uncertainty is reduced to zero. Change in uncertainty is measured by appealing to the entropy of the probability distribution. Entropy refers to the degree of chaos in a system. In the previous example, maximum entropy occurs when each possibility (each day) has the same probability (of being my birthday); i.e. when no information is given. Entropy is reduced when information is conveyed and as a consequence of that the distribution of probabilities changes; see Dretske (1981) for a classical and accessible introduction to information theory.

Given this, the analysis of *mutual information* begins with "the simple idea that the more specific the relationship between a cause variable and an effect variable, the more information we have about the effect after we perform an intervention on the cause" (Griffiths et al., 2015, 532). So, the authors continue, the amount of information that a cause imparts is a consequence of the change in entropy before and after an intervention. Mutual information thus is defined as the change of probability in the effect that arises if we intervene in the cause (Griffiths et al., 2015, 534). Accordingly, when a cause-effect relation shows invariance relations as the interventionist framework posits, mutual information is exchanged between these variables: "any two variables that satisfy the interventionist criterion of causation will manifest some degree of mutual information between interventions and effects" (Griffiths, 2016, 82). With the notion of mutual information at hand, they advance the following proposal:

> SPEC: the specificity of a causal variable is obtained by measuring how much mutual information interventions on that variable carry about the effect variable; (Griffiths et al., 2015, 538)

which essentially boils down to the fact that less entropy in a system entails a higher probability to obtain a certain outcome. Similarly, if the mutual information of two variables reduces entropy, then these variables possess a high degree of biological specificity. In terms of development, a developmental resource that constrains the repertoire of possible outcomes exerts a fine-grained control over the outcome, i.e. it is a source of specificity because it reduces the probability of the outcomes to be obtained. Shortly, in a way reminiscent of Oyama's insight on the role of information in biology, we can say that such specifiers impart order to development. The connection between information and order is well-captured by the informational account of causation. If information is "the cause that imparts order and form to matter" (Oyama, 2000b, 3), information must reduce the chaos (increase of entropy) in a developing system, i.e. it must increment order (decrease of entropy) by reducing the probability of producing a particular outcome.[3]

The connection between mutual information and the interventionist model of causation leads to two insights into the analysis of causation. First, mutual information is asymmetrical, insofar as it presupposes a causal arrow with a specific directionality. In a causal graph, causal relationships are represented with arrows indicating the direction of the force that one item exerts on the other. This does not mean that there could not be bidirectional causation, it just means that the manipulation of C (cause) that entails changes in E (effect) does not imply that manipulation of E would promote changes in C: "our measure of specificity measures the mutual information between interventions on C and the variable E. This is not a symmetrical measure because the fact that interventions on C change E does not imply that interventions on E will change C" (Stotz & Griffiths, 2017, 373). Second, an informational measure of causal specificity helps explain the gradient of control that a cause has in the system. As noted, effects are not necessarily produced by a single cause. Each cause has a degree of specificity that can be measured by appealing to interventions on variables exchanging mutual information: "Thus, the use of this information measure can capture a range of relationships between two variables, from no causal control at all, to fine-grained, highly specific causal control" (Calcott et al., 2020, 247).

It might be relevant to highlight the connection between the probabilistic analysis of information and the idea of distributed specificity. As noted, the degrees of causal specificity of developmental resources are a consequence of the distribution of causes throughout the developmental system. In this sense, having or not having a developmental resource implies a higher or lower probability of developing a particular trait. As explained, this is well captured by appealing to informational

[3] A quite similar probabilistic approach to the notion of biological information was put forward by Scarantino (2015); see Stegmann (2017) for a recent discussion of the notion of information in biology.

theory in the context of the interventionist model. This view of development is also present in other proposals, as, for instance, Gottlieb's probabilistic epigenetics, which rejects a determinist epigenetic model by positing that ontogeny does not rest on a single variable (i.e. informational specificity is not depleted in genes) but on a myriad of developmental resources that define a set of possible outcomes (informational specifiers have a degree of fine-grained control) (Gottlieb, 1997, 2007). Similarly, the idea of cycles of contingency advanced by DST theorists (Oyama et al., 2001) stresses the contingency of development as a matter of the probabilities of outcomes that may result from the matrix of developmental resources. As the reader may have already guessed, this probabilistic view of ontogeny is also connected with Walsh's notion of repertoire, an issue that will be addressed in detail towards the end of this chapter.

8.1.2 Eco-Devo and signaling systems

Ecological developmental biology (henceforth: eco-devo) is a term coined by Scott Gilbert in the early 2000s Gilbert (2001). Gilbert's motivation was the need for a more integrative approach of research in the fields of ecology and phenotypic plasticity, in the wake of earlier insights by Levins and Lewontin (1985) and West and King (1987). This section is mostly based on Sonia Sultan's work, wonderfully systematized in her book *Organism & Environment*.[4] By linking eco-devo's emphasis on signaling systems with the notion of biological information, I will argue that G–S's proposal on biological information is not only theoretically robust but also suitable for different approaches within the DT.

Eco-Devo: an outline

As a first approximation, "ecological development, or 'eco-devo,' seeks to explicitly include the organism's particular environment in studying both the *signaling pathways* and the ecological and fitness consequences of phenotypic expression" (Sultan, 2015, 20; emphasis added).

Eco-devo is therefore an attempt to analyze the context of development and its evolutionary implications. By looking at how organisms perceive their environment through signaling systems, eco-devoists aim to better understand how organisms adaptively regulate their development —i.e. how organisms produce apt traits and self-determine their individual (Darwinian) fitness. As Sultan's work focuses mainly on the ecological development of plants and its evolutionary implications, I will take plants as model organisms that self-regulate their development by means of signaling systems (cf. Gilbert, 2012, for examples of signaling

[4] Also relevant in this context are the works by West-Eberhard (2003) and by Gilbert and Epel (2015); but see also West, King, and White (2003), Casal (2004), and Keller (2014).

systems in animal development). We can appreciate the two basic goals of plant eco-devo in the following quote: "Plant ecological development (eco-devo) aims to firstly, determine precisely how plants perceive and respond to the varying environmental conditions they encounter in the real world and secondly, understand the ecological and evolutionary consequences of environmentally mediated phenotypic outcomes" (Sultan, 2010, 96).

I would like to underline four important issues in connection with the conceptualization of signals. First, to conceptualize signaling systems, we need to distinguish three different kinds of relations in which signaling systems participate:

(i) Cell-environment signals;

(ii) cell-to-cell signals; and

(iii) intracellular signals.

The first category concerns the many sources of environmental cues that provide organisms with capacity of perceiving their external environment —such as biotic, abiotic, or symbiotic cues (Gilbert, 2012). The second category has to do with the internal communicative relations among neighboring cells, i.e. with those cues originating in cells within the organism. Finally, the third category refers to the complex networks of singling pathways within the cell, such as those involved in the regulation of gene expression in development or the regulation of metabolic processes. All of them are important. The core idea of a signal is that of a causal interaction between two elements, in such a way that one of them produces an effect on the other by providing information about a certain state of affairs —be it about the inner milieu of the cell, about other parts of the organism, or about the environment (as in the case of animal communication).

The second point has to do with how signals construct complex signaling systems. Here the aforementioned three kinds of signals get mixed: organism-environment interactions are internally processed by cell-to-cell communications which also produce the activation of intracellular signals. The same may happen in other directions. Signaling systems are multidirectonal and multilevel webs of communication channels where information flows. The picture that arises is that of an organism as a complex network of signaling systems that inform the organism of its inner and outer states.

The third point to note is the physicochemical basis of signals. Surely, more empirical research is needed to understand them adequately, but Sultan already gives us a hint of what it might consist of: "Once perceived by the organism, environmental information is transduced into internal signals by means of chemical and physical cell interactants. Chemical signaling components include pigments and

other receptor molecules; hormones; steroids; metabolites such as sugars, oligosaccharides, fatty acids, aldehydes, alcohols, and proteins; lectins, peptides, amino acids, and nucleic acids; mineral ions; and reactive oxygen species" (Sultan, 2015, 55). The basis of signaling interaction may adopt a different kind of physicochemical substrate and belong to different developmental mechanisms.

The most fundamental point is the fourth one and concerns the fact that signaling systems are central in the adaptive regulation of development. The key idea is that developmental regulation occurs as an adaptive response to the flow of information in signaling systems. This may easily be put in relation with postgenomics, which defends a holistic and systemic view of development, where different resources are central in the specification of developmental outcomes. The principal point of departure with the replicator and gene-eye views of the MS is to treat the cell (an organism) as the central unit of developmental control. It is cells that regulate genes, not the other way around. On this assumption, signals are central in the cellular regulation of development by informing its conditions of existence. In other words, signals enable organisms to produce a response that accords with the context of development. Given that aptness could be understood as something functioning 'according to the context', signals allow organisms produce adaptive outcomes during development, thus "perceiving a relevant first step in any plastic response" Sultan (2004, 231). The adaptive value of signals is a central step toward a teleological view of development. We could articulate this fourth point in another way. If we reject a preformationist view of development that states that developmental outcomes are adaptive because goals are predefined by adaptive evolution, then, what element of development makes developmental outcomes adaptive? The response is the sensitiveness of development achieved by signaling systems that allow organismic regulations to be adaptive to the conditions informed by the signals. As it will be emphasized later, linking signals with biological information will help us understand how signals —taken as causal specifiers— exert fine-grained control in development, and therefore, connect the role of signals in development with the view of developmental control pursued by eco-devo.

Sonia Sultan wraps up the relation between signals and responses under the label of *Environmental cue–response systems*:

> "[T]he pathways that underlie plastic expression patterns can be generally understood as *cue and response systems*: developmental, physiological, or behavioral adjustments (whether adaptive or maladaptive) that occur when an organism *perceives* some aspect of its environment as a specific piece of information and then *responds* to that cue by expressing particular phenotypic effects [...] [T]hese systems comprise signal transduction networks that are embedded in larger regulatory

> networks. Phenotypic outcomes reflect the integration of these webs of environmental cues, their molecular transduction and regulatory pathways, and the feedback that occurs throughout the life cycle" (Sultan, 2015, 49; emphasis in the original).

Environmental cue-response systems exemplify the role of signals in development. The core idea is that adaptiveness lies in the regulation of development due to the information the organism possesses about its inner and outer worlds. Development is adaptive not because it is the result of evolution, but because it is a context-sensitive process, involving both the whole organism and its coupling with the environment. These kinds of input-output or perceptual-response dynamics are well documented in Sultan's work.

Signals and graphs

Let us now turn to outline a proposal by Scott Gilbert and Johnathan Bard on the use of graph theory to formalize the dynamics of developmental processes (Bard, 2011, 2013; Gilbert & Bard, 2014). The core idea of a causal graph is indeed similar to the one of Woodward's interventionist model. Following Gilbert and Bard, a causal graph is defined by a set of triplets. Each triplet is constituted by two nodes being connected by a particular relationship. In Gilbert and Bard's case, nodes may be signals, tissues, processes, cell parts, or any organic entity involved in the dynamics of development. The relations that connect nodes always involve a sort of activity, such as activation, generation, production, constraints, or drive. For instance, a signal can activate a certain genetic regulatory network (GRN) (Triplet 1: signal–activates–GRN). Such GRN can in turn generate a particular process in a specific tissue X (Triplet 2: GRN–generates–process-in-tissue-X), and this tissue may eventually drive a change in a particular morphogenetic structure (Triplet 3: tissue-X–drive–morphological-change); see Bard (2011) for additional examples. In this one, the causal graph is defined by the causal relationships between nodes in Triplet 1, Triplet 2, and Triplet 3. The authors's idea is that any developmental process involves hundreds of triplets and that each triplet must be discovered by experimentation. What graph theory allows then is to appeal to the mathematical tools of graph theory both to deal with the complexity of development (by constructing several sets of triplets) and to picture it in a visual format. This is just a rough sketch of Bard and Gilbert's proposal, but it suffices to highlight their main point: "it turns out that many developmental phenomena can be represented as a graph where the nodes are biological entities scaling from proteins upwards and the edges are relationships" (Gilbert & Bard, 2014, 137).

My motivation for considering their proposal is twofold. First, as the authors themselves note, signals have a central role in the regulation of development: "Like

other developmental activities, the initiation of a morphogenetic process usually results from a signal activating a gene regulatory network that will, in turn, activate the network which drives that process and may also direct the synthesis of some of its proteins" (Bard, 2011, 593). So they picture development as a complex process starting and being coordinated by different signals throughout ontogenesis. In this sense, their account, just as the eco-devo one, also adopts a systems and context-sensitive approach to development.

Now, the introduction of graph theory in eco-devo for the formalization of models of development may be put in relation with the other accounts that also involve graphs —Walsh's analysis of purposive invariance and G–S theory of biological information. Seeing developmental outcomes as the result of many causal relations involving different elements is an idea shared by these proposals. The connection between these proposals will carried out in Section 8.1.4. To see why this connection can be made straightforward, let's see how we can incorporate signals to the informational account of G–S.

Signals as causal specifiers: information and eco-devo

The informational underpinnings of signals can be disclosed by appealing to the G–S framework.[5] Earlier, signals were pictured as those relationships in a developing system that convey information about the inner and the outer world, and such relationships (within the cell, inter-cellular, and with the environment) build complex networks of signaling systems: "Information also flows through [signaling] networks by passing through one node and to the next" (Calcott et al., 2020, 239). What is of interest to us here is that once signals are understood in informational terms, one can apply the causal framework of information constructed by G–S. Calcott et al. (2020, 248) propose "to treat a signaling network as a causal graph, and to measure how causally specific a signal is for an act". This helps, firstly, understand the causal basis of signals and, secondly, treat signals as biological specifiers, which, eventually,"allows us to connect the structure of a signaling network to existing work [Woodward's invariance model] on causal explanation" (Calcott et al., 2020, 244).

This view is a direct application of G–S' proposal to signals. Since signals are defined in terms of casual relationships conveying information, signals are relevant causal specifiers that exert fine-grained control to development. In support of this view, the authors add that

> signalling networks should be treated as causal graphs. This makes explicit the directionality of signalling flows in these networks, and identifies signals as points of intervention, whose manipulation has the

[5]The material in this section closely follows the proposal presented in Calcott et al. (2020).

> power to change acts. Our strategy will be to suggest that the flow of information from a signal to an act should be understood as a causal notion, equivalent to the causal influence that the signal has over the act. (Calcott et al., 2020, 243)

If signals are causal resources in development that promote causal specificity, then the interventionist model can be applied. The grade of control that a signal has may be measured by using the definition presented in Section 8.1.1 on page 243. The result is that the specificity of a signal is obtained by measuring how much mutual information interventions on that variable —the signal— carry about the effect variable. Signals, as information bearers of causal specificity, are therefore central nodes of controls in the causal explanation of development.

Importantly, defining signals in informational terms as causal specifiers sheds light on the role of signals in eco-devo explanations. As commented before, one of the explanatory roles of signals is accounting for the adaptiveness of development. That is, signals endow organisms with context-sensitiveness in such a way that their response is adaptive insofar as it fits with the state of the world informed by the signals. Stotz and Griffiths also attribute this role to signaling systems:

> Signaling arose because the modular structure —the separation of transducer and effector— created a coordination problem. For the organism to respond adaptively, it needed to coordinate these parts, and a signaling system provided the solution. Signaling, from this internal perspective, is a way of building adaptive, plastic organisms. Stotz and Griffiths (2017, 373)

Understanding signals as casual specifiers emphasizes the change from genes to organisms as the units of developmental control that eco-devo supports. As signals are distributed throughout different and complex networks of signaling systems, each signal has a degree of developmental control relative to the degree of specificity it conveys (measured in terms of interventions on mutual information). We can therefore conclude that the organism is the proper unit of developmental control insofar as it is responsible for regulating the different signals —i.e. causal specifiers of development— distributed at different levels of organization.

8.1.3 Biosemiotics

Biosemiotics is a relatively new discipline. As it is usually acknowledged (Kull, 1999, 2001), Jakob von Uexküll is consdered to be one of its main precursors. Biosemiotics is a fairly heterogeneous field in which many different views and proposals coexist, but here I will be mostly based on the so-called Tartu-Copenaghen

School, which is the one with a closest connection with von Uexküll's theoretical biology (cf. "Why Biosemiotics? An Introduction to Our View on the Biology of Life Itself", n.d., for a general introduction to this biosemiotic approach). It was founded by Thomas Sebeok (Sebeok, 1988, 2001), and its main contemporary figures are Joseph Hoffmeyer (Hoffmeyer, 2008a, 2011a, 2011b) and Kalevi Kull (Kull, 2009, 2011, 2021), who will be my main references in this subsection.[6]

The aim of this subsection is mostly programmatic: it certainly does not contribute anything new, but it will help scaffold a number of crucial ideas, in addition to suggest possible bridges between disciplines, such as eco-devo and biosemiotics (Rama, 2021). I would like to do two things here:

1. introduce biosemiotics and argue why it should be part of the DT (i.e. to trace its connections with non-neo-Darwinian views of life); and

2. connect von Uexküll's thought with a central thesis already anticipated: that development is directed towards the U–I, and that the U–I can be understood in terms of signaling systems.

Biosemiotics in the Developmental Turn

A major departure with the MS that makes biosemiotics part of the DT is its rejection of gene-centrism. Hoffmeyer's article "The Central Dogma: a joke that became real" speaks on its own (Hoffmeyer, 2002). The following excerpt taken from this paper directly connects biosemiotics with contemporary post-genomic views:

> Biosemiotics implies a major change in our conception of the role of the genetic material, the genome. Whereas the traditional view sees the developmental process and thus the phenotype as specified by the genetic setup of the organism, the genotype, Biosemiotics makes us consider the overall process as one of interpretation rather than one of specification. (Hoffmeyer, 2008a, 59)

But more important for our purposes here is Hoffmeyer's use of the word *interpretation*. To be sure, this is what characterizes the approach of Hoffmeyer and Kull, and what connects their work with von Uexküll's organicism (see below).

[6] Favareau (2010) and Kull, Emmeche, and Favareau (2011) are two fairly complete compilations of papers including additional relevant material in biosemiotics. As for the other biosemiotic schools, there is Anton Markoš's Biohermeneutics (Markoš, 2002), Howard Pattee's Physical Biosemiotics (Pattee, 2007, 2012), and Marcello Barbieri's Code Biology (Barbieri, 2014, 2015). Although without doubt the work of these scholars is of interest for my purposes, space reasons prevent me from consider it in detail here.

According to Hoffmeyer, organisms are capable of interpreting environmental information. Interpretation should not be taken as involving any kind of robust semantic notion, such as concepts or linguistic capacities typically associated with psychological processes in humans. Rather, interpretation refers here to the fact that the organism receives environmental information that is perceived, processed, and used to produce an adaptive outcome according to the needs and constraints of the whole system. In other words, behind the idea of interpretation lies the very notion of agency; see Arnellos et al. (2010) for the connection between Hoffmeyer's work and the notions of agency, autonomy, ans self-organization. Many scholars within the DT argue, especially since Lewontin's work, and biosemioticians also believe, that organisms are not "passively subjected to universal laws of nature" and see them "as active systems of sign production, sign mediation and sign interpretation, that harness the physical laws in order to live and sometimes to make a more complex living" ("Why Biosemiotics? An Introduction to Our View on the Biology of Life Itself", n.d., 1). The idea of interpretation is therefore tied to the organismal capacity to organize and direct its material conditions to an adaptive outcome. "To act on its own behalf in an environment", as Kauffman (2000, 8) said, the organism must interpret such environment in a way that its action makes sense to itself in such context.

There are many other assumptions that oppose biosemiotics with the MS view of life (Kull, 2021; Noble, 2021), but the central idea of the DT that organisms self-regulate their developmental trajectories, is clearly central to biosemiotics, both through its endorsment a post-genomic view of genetic activity and its emphasis on the capacity of agents to interpret their environment.

With the link between biosemiotics and the DT firmly established, I can now turn to explore the connections between biosemiotics and eco-devo. As argued in Rama (2021), the principal point of contact between the two disciplines is that they both systematically invoke signaling systems in their respective explanations. As already noted, eco-devo appeals to singling systems when explaining the adaptive regulation of development. Biosemiotics, as its name indicates, is devoted to the understanding of signals —the science of semiosis— in living systems (biology). In both cases, the main goal is to comprehend how organisms adaptively respond to environmental information. Eco-devo is centered on development, while biosemiotics also deals with other kinds of biological phenomena, such as animal communication. What they share, therefore, is their way of looking at development as a process guided by signs to produce adaptive outcomes. Note that the phrase *guided by signs to produce adaptive outcomes* already contains a hint that approximates it to intentional explanations. I'll come back to this later.

Understanding the U–I as informational sates

In addition to the aforementioned observations and the pertinence of taking biosemiotics as part of the DT, there is another important element in biosemiotics that makes it relevant both for the refurbishing of biological theory and for the purposes of this chapter. This concerns the influence of von Uexküll on biosemiotics; see Kull (1999, 2001) for a historical analysis. Once the central role of signaling systems in adaptive developmental regulation is recognized and signals are defined in informational terms as causal specifiers, we can establish a link with an insight taken from von Uexküll's theoretical biology, namely the idea that goals are not related to adult phenotypic states but with a harmonic regulation of the U–I. The core idea is that both biosemiotics and eco-devo enable a picture of the U–I in terms of signaling systems and that this is central to an account of the directedness of development.

The connection between signaling systems and the U–I is relatively straightforward to establish, because after all biosemiotics owes to the biological thought of von Uexküll a large part of its theoretical underpinnings. The interpretation effected by an organism is what constructs its U–I. In the biosemiotic jargon, those intrinsic processes that lead to the construction of the Innenwelt are labeled 'endosemiotic processes', whereas the Umwelt —or the 'inside exterior' (Hoffmeyer, 1998, 33)— is constructed through 'exosemiotic processes'. Hoffmeyer (2011a) introduces a clarifying qualification in this distinction with the notion of 'semiotic niche' as the 'outside interior' (Hoffmeyer, 1998, 33), which helps distinguish the Umwelt —to refer to signaling systems bearing information about the environment— and the semiotic niches —to refer to those elements of the world that are experienced by the organism. The dichotomy between Umwelt and semiotic niche parallels the distinction between sign and reference. The Umwelt is the inner signaling state informing about the environment, while the semiotic niche is defined as those external elements of the environment that the Umwelt refers to.

The idea of von Uexküll's U–I can also be found in eco-devo albeit not in an overtly explicit form. This was already observed in connection with Denis Walsh's account of teleological development. Specifically, the idea that organisms experience the environment in a particular way in order to provide an adaptive response was already connected with von Uexküll's work in Section 7.3. However, within eco-devo we can find a more direct pathway to this idea in Sonia Sultan's work. As noted in Section 4.1, niche construction theory has different interpretations depending on the kinds of phenomena it is applied to. While standard views concern populational processes (Odling-Smee et al., 2003), Sultan's provides an individual, developmental view of niche construction. Moreover, the notion she has in mind does not refer to the material changes that the organism exerts in the external environment, but she rather is interested in how an organism experiences

its niche. She puts a lot of emphasis on what could be named *experience niche construction* (Section 4.1.2), that is the particular way in which each organism becomes informed of the environment (Sultan, 2015). The ideas of experiencing and interpreting the environment takes us back to von Uexküll's place. It is almost inevitable not to perceive the family resemblance between von Uexküll's account and Lewontin's rejection of Explanatory Externalism. Compare this pronouncement by Kalevi Kull, "[t]he main question of biosemiotic methodology is: *how to know what organisms know*" (Kull, 2009, 83; emphasis in the original), with Lewontin's claim that "if one wants to know what the environment of an organism is, one must ask the organism" (Lewontin, 2000, 54).

So once we have traced the parallelism between the U–I and signaling systems, the next step involves understanding U–I as informational states. This is also quite straightforward. Given that signals are defined as causal specifiers providing biological information, a network of signaling systems provides information about the state of affairs of the U–I. We can now spell out von Uexküll's idea that goals are directed towards a harmonic relation with the U–I. The idea is pretty simple: organisms construct the U–I through signaling systems and their responses are directed to adaptively (i.e. harmonically) regulate the system according to the U–I.

As I already advanced above, this is the core of my strategy to deal with teleological explanations in development as intentional explanations of sorts. Informational states causing an effect precede the goal that such an effect fulfills. Having a goal can then be defined in informational terms. In development, signaling networks causing a developmental response precede the goal that the developmental response fulfills. This idea will be developed in detail in the next section.

8.1.4 Causal graphs: the shared space

The aim of this last subsection is to put together the different ideas presented so far. In part, this task has been already accomplished in earlier sections. On the one hand, I underscored the connection between the informational account of G–S with both eco-devo and biosemiotics and, on the other hand, I also argued that a close relation exists between eco-devo and biosemiotics. As I see it, these three approaches point to the same direction and contribute in their own way to the pursuit of a non-gene-centered view of development. G–S's proposal is a robust theoretical account of the distribution of causes supported by the post-genomic framework, eco-devo is mostly an experimental research area pursuing a holistic and context-sensitive view of development, while biosemiotics contributes to the whole with important philosophical insights.

As a first step towards the proposed unification, I will focus on an element shared by the different accounts presented here: the interventionist model of causal

explanations and the use of causal graphs to model developmental dynamics. This idea is present in Walsh's Agential Perspective, in the G–S's definition of biological information, in eco-devo's modeling of development, and is also central to understanding the causal basis of teleological explanations. Figure 8.1 shows an informal casual graph. The common ground of the theories presented so far lies in the fact that each proposal can be framed in terms of something similar to Figure 8.1. I will analyze each proposal in turn by attributing different interpretations to the variables X_n and the outcome Y.

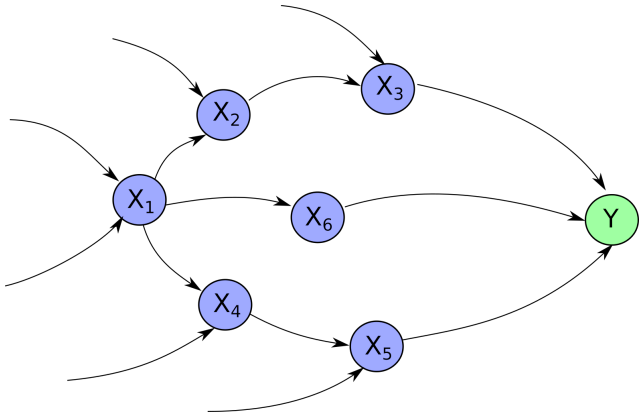

Figure 8.1: An informal causal graph.

Woodward's causal analysis

Woodward's causal analysis interprets generically Xs as causes (C_n) and Y as the effect (E) (Woodward, 2002, 2003, 2010). C_ns are determined on intervention. Variations on C_n must result in a variation on E. An explanation is causal if it involves variables that exhibit invariance relations. An assumption that Stotz and Griffiths (2017, 372), citing Woodward, also endorse: "Causation is conceived as a relation between variables in an organized system that can be represented by a directed graph. A variable X is a cause of variable Y when a suitably isolated manipulation of X would change Y". Figure 8.2 on the following page illustrates what would be the classical view of causal relationships modeled with causal graphs as developed by Woodward in Chapter 2 of his 2003 book, and strongly influenced by the philosophical views of Wesley Salmon (Salmon, 1984, 1998) and the formal work of computer scientist Judea Pearl (Pearl, 2009),[7] among others.

[7]For an accessible introduction to the formal approach to causality developed by Pearl, see Pearl and Mackenzie (2018).

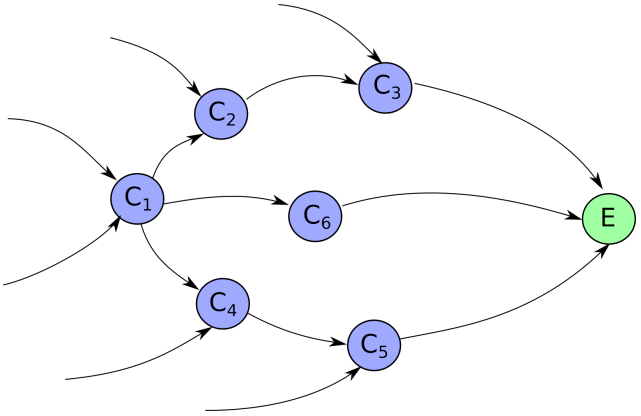

Figure 8.2: Causal explanations in the interventionist framework.

G–S causal analysis

G–S analysis interprets Xs as causal specifiers (CS_n) conveying biological information and Y as a developmental outcome (DO). As explained, causal specifiers are defined on the basis of mutual information. They exhibit different degrees of developmental control: the amount of information that a causal specifier provides is the result of the variation in entropy that results in DO by intervening on CS_n. If entropy decreases, a rich amount of mutual information is transmitted from CS_n to DO, i.e. CS_n is a relevant source of biological specificity in development. This is captured by Figure 8.3:

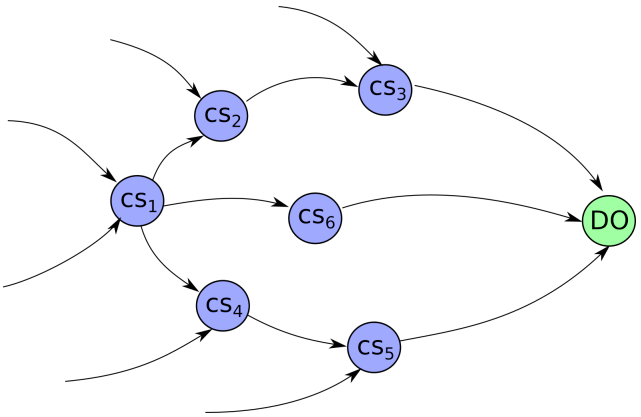

Figure 8.3: Causation and biological information.

Eco-Devo's causal analysis

While Gilbert and Bard do not explicitly endorse an interventionist account of developmental dynamics, they do appeal to causal graphs as a useful tool, albeit with some limitations (Gilbert & Bard, 2014, 136), to model developmental processes. If we focus on the role of signals as causal specifiers, the Xs in Figure 8.1 on page 255 should be interpreted as signals (S_n) and Y as the developmental outcome (DO) these signals promote. Under this interpretation, we get Figure 8.4:

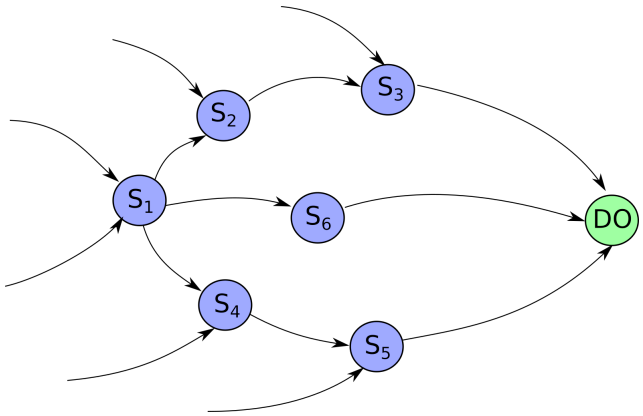

Figure 8.4: Signals as causal specifiers.

Walsh's causal analysis

Having shown the equivalence of these different analyses of causation by reducing them to causal-graph models, we can now move to Walsh's deployment of the interventionist framework. Here, it is important to keep in mind the criticisms presented in Section 7.3.3. According to Walsh's notion of purposive invariance, relations are between goals and means. However, recall that we cannot interpret Xs as goals (G) and Y as means (M), because goals —as end-states— are the result of a developmental outcome, not the other way around. They do not take place before developmental responses. This is the very problem of teleological explanations: How future consequences can have a causal influence on current activities?

Alternatively, I argued that invariance relations are established between a system having a goal and means: if organisms are goal-directed to a different developmental outcome, a different developmental pathway may be chosen. Therefore, Xs should not be treated as the goals of development, but rather Xs should refer to a state of an organism having a goal, while Y is a means (M) toward the achievement of the goal. Importantly, as I will argue in detail in Section 8.3, having a goal is not a property of a single causal specifier or of biological signals.

Rather, having a goal is a global and emergent property of developing agents that integrates different biological information in order to produce a particular means. Each Xs, therefore, refers to any source of biological information (BI_n) —any developmental resource conveying causal specificity. With these amendments, the corresponding causal graph would look as depicted in Figure 8.5.

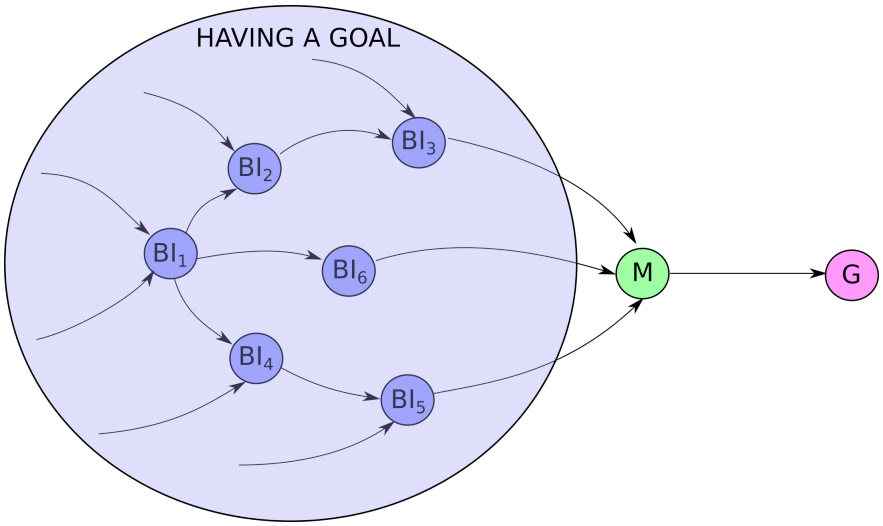

Figure 8.5: Causation in teleological explanation.

A pending question

In this subsection, I shall carry out my attempt at unifying the different approaches discussed through a reduction to causal-graph models. This move is essential in order to improve our understanding of teleological explanations, because, as argued, understanding causal specificity in informational terms makes it possible to ascribe causal roles to informational states in developmental control. Such a causal roles constitute the causal land where teleological explanations are rooted.

Before moving on, however, a pending question needs to be settled: How does this causal analysis of biological information lead to an intentional view of information? It is important to recall from our discussion in section Section 3.3.2 that two different notions of information coexist in the life sciences. One is the *causal or factive* notion of information, and the other is the *intentional or semantic* notion of information. In G–S's proposal, the invoked notion is the causal one: information is defined in terms of the causal connection between different elements in a system and the analysis of informational specificity is based on interventions in a causal graph, and Crick information —extended beyond the gene— is formalized by appealing to Shannon's mathematical theory. G–S's analysis of biological specificity, as they note, is just about causal specifiers in the living realm.

The question as to whether intentional information should be also posited remains open, though. Earlier, I pointed out that some etiological accounts of development treat information in intentional terms. John Maynard-Smith, for example, takes natural selection as the source of intentionality that is found in developmental explanations that involve genetic information: "the concept of information is used in biology only for causes that have the property of intentionality [...] A DNA molecule has a particular sequence because it specifies a particular protein, but a cloud is not black because it predicts rain. This element of intentionality comes from natural selection" Maynard Smith (2000, 189-190). Etiological accounts, however, are not necessarily all gene-centered. Shea's theory of inherited representations, for instance, accepts both extended inheritance systems and the Parity Thesis. We have already discussed the problems with this view as a consequence of its tight link with a replicator stance (5.2.3). So the question about intentional information needs to be further constrained: is it possible to come up with an intentional notion of information based on individual-level causes? The same question arises once we define signals in informational terms. One may say that a purely physicochemical account of signals is enough to understand the role of signaling systems in development and that such an account would not be teleological. Note that this kind of arguments are often also wielded within cognitive science: one can in principle provide a purely neurophysiological explanation of behavior in terms of electrical connections in networks of neurons. The problem with these explanations is that they block a treatment of behavior as a goal-directed phenomenon and reduce it to a mere output of the system.

Note that this question targets the very necessity of teleological explanations. If development is just a matter of the interaction of different developmental resources, then mechanistic explanations would be enough, teleofunctions would not be necessary, and Cummins-functions would do all the work. Once we look into the causal structure of teleological explanations and argue for the need of some sort of intentionality, then automatically normativity enters the scene: normativity comes to the fore at the limits of mechanicism, when a purely physicochemical mechanical account falls short in the explanation of development.

I can now try to link this question with the analysis of the causal structure of teleological explanations. In the previous section, I sketched the relationship between having a goal and means in an informal causal graph —Figure 8.5 on the preceding page. I argued that such a relation is invariant. My appeal to intentional explanations seeks to define having a goal (HG) in *intentional terms*. HG is an intentional state that produces means to take a particular form. My proposal, therefore, accounts for the temporal dimension of teleological explanations of development: goals do not precede means, but representations of goals do precede means. That is the reason why I see this as a sort of intentional expla-

nation. So far however, I have only the tools for arguing that having a goal can be defined in informational terms, not in representational ones. This is, indeed, the difference between an intentional and a purely causal notion of information. Figure 8.6 illustrates this situation. I have argued that the *causes of teleological explanations* concern the relation between a system having a goal (HG) and the means toward the fulfillment of a goal (Figure 8.5 on page 258). Moreover, the proposal of G–S enables treating HG in informational terms. The *informational causes of teleological explanations* are relations between informational states (IS) and means. The pending question concerns understanding informational states in intentional terms. This would lead to positing *representational causes of teleological explanations*: representational states (RS) as causes of means that are directed towards the achievement of goals.

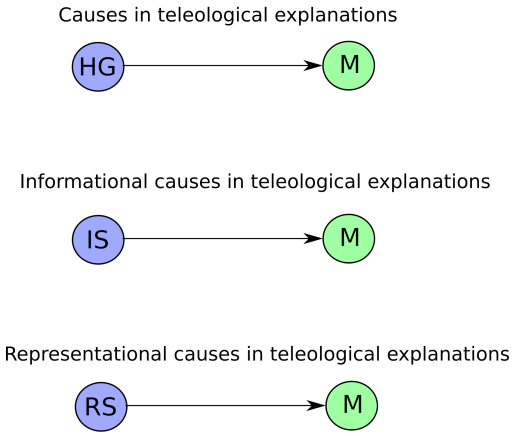

Figure 8.6: The transition towards intentional explanations.

Now, is there intentional information in development? Do signals have content? These questions are not new in philosophy. Most views on signals and information trace back to the work by David Lewis on signaling games (Lewis, 1969) and by Fred Dretske on information in cognitive systems Dretske (1981). In the contemporary literature, there is an ongoing debate about whether and how a contentful view of signals and information can be defended in biology; see Artiga (2014a); Birch (2014); Frick, Bich, and Moreno (2019); Ganson (2018); Godfrey-Smith (2014); Martínez (2015); Shea (2018); Skyrms (2010), and the papers in Birch, Martínez, and Artiga (2020). Some of these works present analyses of whether errors can arise in signaling games, while others ask whether some sort of functional account of information may result in intentional information. Signaling games and sender-receiver systems are the main models within which the nature of intentional information is discussed. Although I shall borrow some ideas

from these authors, I will not enter in a detailed analysis of their models. This is not to be taken in the sense that my proposal deviates from the path initiated by Lewis and Dretske. Quite the contrary, Agential Teleosemantics is firmly aligned with such a path and although it falls out of the scope of this thesis, the approach developed in this chapter should be connected with models of signaling games and sender-receiver systems.

The track towards intentional information I wish to follow now should be quite expected: to defend a notion of agency in development that introduces a normative dimension into the dynamics of development portrayed by G–S. Note that the fact that developing organisms need not be autonomous organisms in all developmental stages does not mean that the philosophical foundations of organicism are irrelevant to development. In other words, promoting a view of developmental agency motivated by developmental organicism would lead us to crucial emergent properties, such as normativity and, eventually, intentionality. Organicism is central to moving beyond purely mechanistic explanations of development, and to arguing that teleology is required.

8.2 Interpreting information

In may quest for teleological development, I argued that development proceeds through informational processes. It is therefore necessary to account for the idea how such information processing requires some sort of intentionality. This is the aim of the present section. To do so, it would be necessary to define, at the developmental scale and in connection with biological information, the central elements of Walsh's proposal: developmental goals, developmental agents, developmental repertoires, and the experienced environment.

8.2.1 Goals and time

I will first introduce the notion of a *developmental stage* which helps appreciate the temporal dimension of developmental goals. Next, I will introduce what according to my proposal developmental goals amount to.

Developmental stages

In such classic textbooks as Gilbert and Barresi (2010), developmental stages are defined as those points in the sequence of developmental processes where the developmental system exhibits relevant changes in relation to the previous stage and the following one. For example, the typical developmental stages of animal development are the zygotic stage, cleavage, blastulation, gastrulation, and so on

(including also cognitive development). At each stage, we observe relevant differences in relation to the previous and next stages. Moreover, the classification of developmental stages can zoom in and zoom out depending on the level of analysis. In this sense, for instance, neurulation, in organogenesis, can be subdivided into two further stages (primary and secondary neurulation). However, if we reach a high degree of granularity, stages may temporarily and spatially overlap. For instance, primary neurulation can be divided into four additional stages, elongation and folding of the neural plate, bending of the neural plate, convergence of the neural folds, and closure of the neural tube, although these stages are not really separated in time and space but rather overlap with each other (Gilbert & Barresi, 2010, ch. 13).

In this thesis, I will adopt a slightly different view of developmental (or ontogenetic) stage. My proposal is exclusively based on time and is purely methodological: a developmental stage is *a specific point in time in development*. Accordingly, there cannot be temporal overlapping developmental stages. Note that my aim is *not* to argue that my conception of developmental stage better captures the facts of development than the concept of stage used in developmental biology as described above. I might have chosen another label, but I decided to use the expression of *developmental stage* because it works well for illustrating the relevant temporal dimension of teleological explanations in development in opposition to adult-oriented views.

Developmental goals

Throughout this thesis, I have advanced a number of observations about the nature of developmental goals that have driven me into a difficult scenario. I cannot adopt an evolutionary view of goals and say that goals are defined by inherited information. Nor can I say that goals are directed to phenotypes in the adult period. Moreover, the goals of a developing organism must be different from that of self-maintenance, even though developmental systems must self-maintain themselves while developing. My view must accommodate all these points and explain the chaining nature of developing organisms. What feature may we identify that is common to all developmental stages?

We may say that the goal of a developmental system is to *change*. But of course, this characterization would be applicable to any system that changes over time, including non-living systems. It might be more adequate to say that the main goal of a developmental system is that of *adaptive change*. This certainly introduces the distinctiveness of living organization and ties teleology to the explanation of aptness, as it should be. However, it is also necessary to identify what changes during developmental processes; i.e. what is modified in the organism during development. What changes during developmental processes is the

constitutive dimension of the organism —its parts, the relation between them and with the environment (cf. below for further comments). Finally, to complete my proposal, it is necessary to emphasize the temporal dimension of developmental goals, i.e. the fact that developmental goals are relative to ontogenetic stages. I conclude, therefore, that the goal of a developing organism is that of *producing an adaptive change in its constitutive dimension at each developmental stage*. Such change makes the developing organism move to the next developmental stage. Development proceeds by the continuity of adaptively directed changes throughout its developmental stages.

Figure 8.7 sums up the core idea by highlighting the temporality of developmental goals (ontogenetic stages) and the sort of change developmental goal-directed responses promote (changes in the constitutive dimension of the organism). Concerning the last point, the organicist tradition and other related frameworks —that gave birth to autonomous systems theory— provided the key to understand the constitutiveness of living systems (cf. Chapter 6). The constitutive dimension of a system refers to those parts that interact one with the others and with the environment and assume a particular causal role that contributes to preserving the system —what later on was known as operational closure. The idea, therefore, is that development is the process that builds the constitutive dimension of living systems in an adaptive way by promoting adaptive changes at each ontogenetic stage. As the system changes through developmental processes the constitutive dimension also changes and new structures and functions are constructed. Directedness is responsible for producing and organizing matter. Such a process is carried out through a continuum of adaptively directed responses acting on the organism itself (and its coupling with the environment).

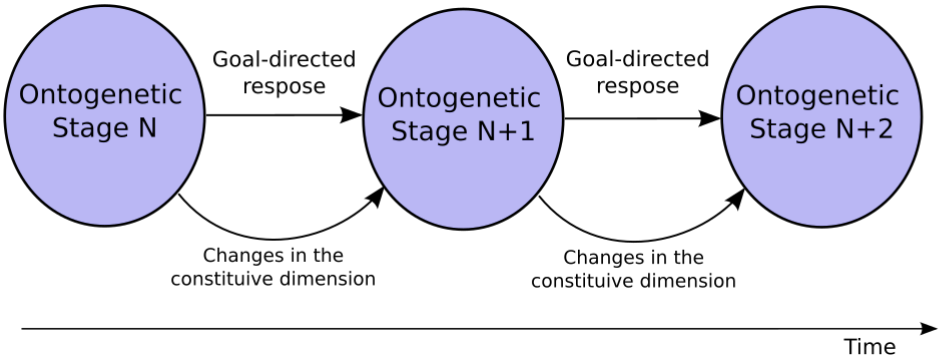

Figure 8.7: Goal-directed changes during ontogenetic stages.

The relation between both ways of being

The view presented before also suggests a picture of the relation between both ways of being. As noted, both ways of being, one related to change, the other to maintenance, overlap in the lifespan of an individual. Both change and maintenance are *always* goals of the system. My point now is that developmental processes are responsible for defining the system that needs to be self-maintained. By changing the constitutive dimension of the organism, developmental processes establish the norms and goals that physiological processes obey. Teleological development (adaptive changes in the constitutive dimension at each developmental stage) in a nutshell determines teleological physiology (self-maintenance). The causal role of a trait in a physiological process that contributes to self-maintenance is determined by its ontogenetic history, that is by such a process responsible for organizing matter, creating order, and generating apt systems.

In Section 6.3.3, I introduced two important notions related to normativity: *norm-establishment* and *norm-following*. The former is the process that establishes a norm, and the latter is the process of acting according to a norm. In teleological physiology, norms are established according to the current organization of the system and their environmental conditions (i.e. physiological constraints), while norm-following concerns those (physiological) activities that contribute to maintaining such organization in viable conditions. I will argue that the situation is similar in teleological development. Norms are established according to the biological information (i.e. developmental constraints) at a specific ontogenetic stage. Norm-following concerns how the developmental system uses its biological information in order to provide an adaptive developmental outcome.

It is relevant also to emphasize the relation that exists between norms in development and in physiology. I already argued that teleological development determines teleological physiology; i.e. that developmental processes construct the constitutive dimension of living systems. Such constitutive dimension defines, at each stage in the lifespan, the norms that physiological processes must obey. Developmental processes are responsible for the construction of the organization that establishes the norms of developmental processes.

8.2.2 Limited by conditions...

One of the central elements of goal-directedness is the fact that a system may have different outputs, and that the outcome that is finally produced is the result of the system's regulation of development. The viability space in autonomous systems theory or the design space in the MS also connects directedness with the capacity of selecting one possibility among many. I will therefore move now to define the repertoires of development which, as should be expected, are relative to

each ontogenetic stage.

Biological information as a developmental constraint

Developmental constraints have been a central topic in evo-devo since its inception. In this context, since the pioneering work of Pere Alberch (Alberch, 1980, 1991), developmental constraints are seen as imposing limits on possible developmental outcomes. Developmental constraints may be of different kinds, such as DNA, environmental conditions, or the physical composition of developing systems. What developmental constraints evidence is that not all variation is possible. I prefer however to look at this from the opposite perspective and express exactly the same by saying that only some kinds of variation are possible. In this sense, a developmental constraint allows certain possible outcomes while forbids others. A developmental constraint *specifies* the limits of what can be achieved in development.

These introductory ideas to developmental constraints in evo-devo can be integrated into the informational framework of the previous section. Any developmental resource that conveys causal specificity to development may count as biological information. A causal specifier provides fine-grained control of developmental outcomes: each specifier allows for a range of possible outcomes, as illustrated in Figure 8.8. As noted, once we endorse a distributed specificity view of development, the mapping between biological information and developmental outcomes is not a one-to-one mapping but a one-to-many relation. Note, therefore, that biological information operates as a constraint on development. The biological information of a developing system specifies the limits of what can be achieved in development. I propose, therefore, that *biological information in development defines the constraints of the developing system.*

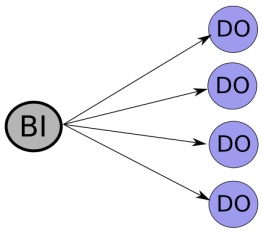

Figure 8.8: Constrains and biological information: biological information (BI) constrains the possible developmental outcomes (DO). Biological information does not determine the outcome, nor does it enable any possible outcome. It is a one-to-many relation.

Biological information and developmental repertoires

The idea of developmental constraints as limiting developmental outcomes was already implicit in the first part of the fifth feature of goal-directedness proposed by Russell (E. S. Russell, 1945, 110; cf. Section 7.2.3): "Goal-directed activity is *limited by conditions*", and being limited by conditions is being constrained by developmental resources (i.e. biological information). As in Russell's proposal, biological information ('the conditions of development') defines the *repertoire* of possible outcomes.

My proposal is, indeed, similar to the ones by Russell and Alberch. Development has a repertoire of possibilities. While in my view such repertoire is defined in terms of biological information, Alberch appeals to 'developmental constraints' and Russell to 'the conditions of development'. The main and most relevant difference between my proposal and the others is that my view is explicitly non-adult-oriented, while the other proposals represent a repertoire of *adult* outcomes. Of course, there is no problem with providing an adult-oriented view of the repertoire if one is really needed. Indeed, biological information can also provide an adult-oriented repertoire. But my point is that no reference to adulthood is needed for my view of teleological development. In this sense, to make my exposition clearer, I shall use the expression *developmental repertoire* to refer to the repertoire of possible responses at *each* developmental stage. As a result, the repertoire of possibilities that is relevant for teleological explanations of development concerns the *next ontogenetic stage*. This idea comes from my view of developmental goals and their temporal dimension.

Developmental goals are relative to each ontogenetic stage: biological constraints at ontogenetic stage N define a repertoire of possibilities at ontogenetic stage $N + 1$. Aligned with this view, the notion of *developmental outcome* does not point in the direction of adulthood. Rather, a *developmental outcome* is relative to the *next* ontogenetic stage: a *developmental repertoire* defines the possible *developmental outcomes* that may be achieved at the next ontogenetic scale.

Once these terminological clarifications are made (motivated by the temporal dimension of teleological explanations), the idea is that biological information constrains what can come next. At each developmental stage, there is a specific configuration of causal specifiers, and consequently a particular developmental repertoire. A developmental outcome of stage N locates the developing system in stage $N + 1$. This idea is presented in Figure 8.9.

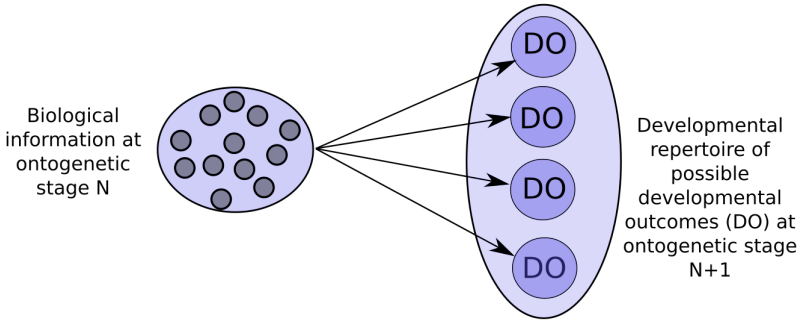

Figure 8.9: Biological information and developmental repertoires: black dots represent biological information. The set of pieces of biological information at ontogenetic stage N defines the developmental repertoire of possible outcomes at stage $N + 1$.

Moreover, we can see that each developmental outcome in the repertoire produces a particular configuration of biological information. Biological information at stage N defines a repertoire of developmental outcomes at stage $N + 1$. Each possible developmental outcome in the repertoire has a particular set of causal specifiers. The are two reasons why causal specifiers are modified during development. First, many environmental inputs are absent at some stages but arise in future ones. For instance, extended systems of inheritance, particularly those that provide environmental information, are not depleted at conception. Quite the contrary, environmental inputs participate at different stages, in such a way that new causal specificity may arise at certain points of development. Second, many developmental resources are constructed by the developing systems themselves: causal specifiers present at a particular stage may be generated by the developing system at previous ontogenetic stages (in Section 8.3.1 I will say more about the different sources of biological information). In sum, the core idea is that depending on the trajectory taken by the developing system, different biological information may take part in development. This situation repeats stage after stage: each possible developmental outcome at stage $N + 1$ defines a repertoire for stage $N + 2$ and so on and so forth. This view captures the changing nature of development. Development moves on by choosing a specific developmental outcome at each developmental stage. While development moves on the structure and functioning of the organism starts to emerge and new repertoires of developmental outcomes arise at each stage. This situation is illustrated in Figure 8.10.

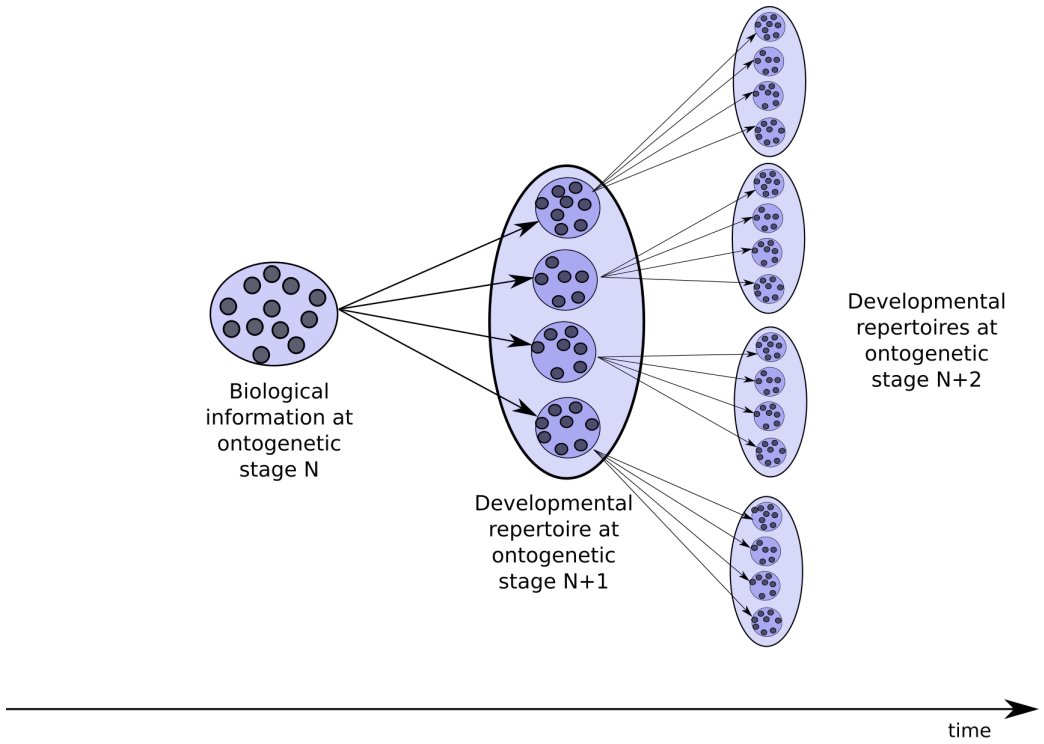

Figure 8.10: Constraints and repertoires throughout ontogenetic stages: black dots represent pieces of biological information. Biological information at a particular ontogenetic stage N defines the repertoire of outcomes at $N + 1$. Each outcome is itself constituted by biological information, which defines further repertoires at the future developmental stage $N + 2$.

Note the similarity between constraints and repertoires in the developmental scale and constraints and repertoires in the physiological scale. In autonomous systems, the operational closure of constraints (coupled with the environmental conditions) defines a space of possible physiological activities, just as developmental constraints define a repertoire of possible developmental outcomes. However, in an autonomous system, we already saw that the repertoire of possible physiological activities has an adaptive value: the viability field (Section 6.4.3). In this sense, within what is possible, an autonomous system would choose what is necessary: whatever that contributes to self-maintenance —i.e. what contributes to fulfilling the goal. However, in development, "adaptive values are not the evidential basis from which the constraints are inferred" (Amundson, 1994, 564). In other words, as it is clear from the structuralist underpinnings of evo-devo, a repertoire of possibilities does not make adaptive valuations. That is, it says nothing about which possibility is adaptive and which one is not. This requires connecting biological in-

formation and developmental repertoires with developmental agency. This comes next.

8.2.3 ...but not determined by them

To get to intentionality, we need the second part of Russell's quote: "goal-directed activity is limited by conditions *but not determined by them*". Given that the conditions of development were identified as the developmental resources that convey biological information, this quote defines the limit of mechanicism. Development, as a goal-directed activity, is not determined by its developmental resources; i.e. it is not determined by the interaction of the parts of a developmental system. It is at this point that key emergent properties, such as agency, norms and, eventually, intentionality enter the scene.

A gradual and multi-level view of developing agency

In Chapter 6, following Barandiaran et al. (2009) I presented three central defining properties of agency —individuality, interactional asymmetry, and normativity. I also showed that autonomous systems are agentive systems. The issue now is to see whether the three defining properties of agency are present in development or not. Are developing systems agentive systems?

Following the analysis of Nuño de la Rosa (2010) presented in Section 7.3.2, a first difficulty arises once we note that it is not easy to identify in all developmental stages some of the defining properties of agency in the whole developing system. For instance, at some developmental stages of vertebrates, such as blastulation, there is no clear boundary between the system and the environment, in such a way that the very idea of interactional asymmetry does not make full sense. Also, at some other stages, for instance during gastrulation, the system can be divided into two different ones, in such a way that two distinct organisms are developed. At these stages, following the etymological definition of individuality, the organism is not an in-dividual (i.e. it can be divided). Besides, while the zygote may be taken as an agentive system (cells were identified as minimal agents; Section 6.3.2), agentivity does not seem to be present at the very first stages of development at the level of the whole organism, but they rather appear once the system has achieved a certain degree of complexity. My treatment of agency must take these considerations into account.

As for the connections between development and agency, I would therefore like to suggest that the two following properties hold:

(i) Agency *gradually emerges* at the level of *the whole developing system* during ontogeny.

(ii) Agency is *relative to the level of organization* achieved in a developmental process.

Concerning (i), the idea is that the three defining properties of agency become gradually manifested as development unfolds. First, a *boundary* between the environment and the organism is created as development goes on. Second, the capacity of *modulating* environmental inputs arises while the system is differentiated from the environment. Third, the adaptive modulation (i.e. regulation) of developmental processes requires both that the system is differentiated from the environment and that the system is capable of integrating environmental resources to provide an adaptive response (more on normativity later on). Note that here I am discussing whether the *whole developing system* should be considered as an agent or not. Later I will address the issue of whether its parts should be treated in agentive terms, but when we focus on the whole developing system we realize that agentive properties are not present at all ontogenetic stages but that they gradually emerge as the process goes on.

Claiming that the developing system acquires agentive capacities somehow implies that the whole developmental system becomes central to explaining its own developmental processes. That is, the idea introduced in Section 4.1.1 stating that organisms are the proper unit of analysis in developmental processes—not their genes or any sub-organismal part. When the developing system acquires an agentive character, the interaction between the different parts of the system starts to play a crucial role: we observe an integrated system where its parts interact —for instance, through signaling communication— to produce an outcome. Similarly, environmental inputs start to play a crucial explanatory role in the control of development and the activation of different developmental pathways. The system becomes capable of perceiving an environmental signal, processing it through different signaling pathways, and producing an adaptive response. In sum, the emergence of agentivity in the whole developmental system implies that the whole system participates in the orchestration of its own developmental processes, both by integrating environmental biological information and regulating its internal dynamics.

As for (ii), agentivity is also relative to the level of organization. If we look at a particular developmental stage —e.g. gastrulation— maybe we find out that the developing system as a whole is not an agent, but, if we zoom in, we will find that it is made up of different elements that deserve to be treated in agentive terms: cells that differentiate and reproduce. Agency is therefore present in certain parts of the whole system. This is just a consequence of the fact that an agentive system may be composed of other agentive systems. Since cells are typically considered to be the minimal case of an agentive system, we should always understand cellular activity throughout development in agentive terms, and accordingly, explain

cell development in teleological terms. While this fact does not sanction a treatment of the whole developing system in teleological terms —this comes gradually in ontogeny, it is a clear indication of the presence of agentive properties at *all* developmental stages.

This gradual and multi-level view of agency is operationally useful for dealing with teleological development, while at the same time it acknowledges the fact that not all the defining properties of agentivity are present at all developmental stages *at the level of the whole organism*.

Self-regulation: interpreting and using information

In this third part of the thesis, the key notion is agentivity. The centrality of organismal agency is antagonistic to the view of life where genes are the units of control and organization. In Chapter 6 I observed that to the extent that autonomous systems are agents, they can *regulate* physiological processes in a teleological way. The same happens at the developmental scale. Development is a teleological process not because it obeys a genetic program but because it is a process *regulated* by an agent. The main idea around agency is that of *adaptive self-regulation*: developmental goals are achieved because *adaptive changes at each ontogenetic stage are harmonically orchestrated* —to use von Baer's musical metaphor— *by agents*.

Like other central phenomena in teleological explanations of development, self-regulation can be understood in informational terms. A developmental system self-regulates its ontogenetic trajectory by interpreting and using its biological information to produce an adaptive outcome. Regulation is about using means to achieve a goal: to use developmental resources in a way that an adaptive output is obtained.

Here, we can enrich the idea of self-regulation as the interpretation and use of biological information by appealing to the theory of information itself. The main idea around interpretation is that, in order to provide an adaptive response, the systems must integrate different sources of biological information, it has to process information coming from different developmental sources. In informational theory, the integration of different sources of information gives rise to what is known as synergistic information (de Llanza Varona, 2022; Griffith & Koch, 2014; P. L. Williams & Beer, 2010): interpreting biological information cannot be accounted for just by taking into account each specific source of information independently but, rather, it is necessary to take into account the whole matrix of biological information. A suitable illustrative example can be found in cognitive science concerning multi-modal perception (de Llanza Varona, 2022). Multi-modal perception concerns those perceptual representations that arise by integrating information from different perceptual modalities (e.g. visual, auditory, etc.). In those cases, to define the content of a representation, it is not enough to look at the in-

formation provided by a single modality. Instead, the information from different sources is integrated to achieve a particular perceptual representation. I contend that the distributed specificity of developmental resources suggests a multi-modal view of biological information in development. This does not mean that the different modalities that provide biological information are perceptual modalities; it just entails that interpreting biological information is about processing and integrating pieces of biological information coming from different sources.

This view fits with my claim that the causal structure of teleological explanations in development is identical to the causal structure of intentional explanations. As argued in Section 8.1.4, having a goal can be understood as a representational state composed of different sources of biological information that an agent processes in order to provide an adaptive outcome. The fact that agents pursue goals and that this process can be described in informational terms introduces the two core elements of intentionality: aboutness (due to informational states) and normativity (due to goal-directedness). In this sense, when I state that interpretation is about integrating different modalities of biological information, this means that the result of such integration is a representational state. As in the case of multimodal perception in cognitive systems, the representational content of an agent interpreting biological information cannot be reduced to any specific source of biological information. From this viewpoint, we can understand self-regulation in the context of Russell's quote: to produce an adaptive outcome within the limits imposed by its material condition; or, in informational terms, we can understand self-regulation as the interpretation and use of biological information: processing information to generate a representational state that conduces the system to achieve its goals.

Normativity

As noted throughout this thesis, teleology and normativity are interconnected: goals define the norms that a system must obey in order to achieve the goal. The distinction between 'the double way of being' of organisms motivated the hypothesis that there are two distinct kinds of goals: goals about change —*the changing way of being*— and goals about maintenance —*the maintenance way of being*. As norms are defined by goals, *developmental norms should be different from those norms connected with self-maintenance*. Developmental norms concern the adaptivity of a developmental change. A developmental process must do whatever is necessary to produce an adaptive developmental response within the developmental repertoire. If this process does not result in an adaptive response, the developmental system would not fulfill its goal. If the organism at a particular developmental stage, say organogenesis, is directed to the production of a particular organ, and for some reason (to be explained below) this does not take

place, a maladaptive outcome may result. Normativity in development, therefore, is related to the maladaptiveness and adaptiveness of phenotypic outcomes.

A central guiding idea in this thesis is that the responsibility of producing traits cannot be subsumed by any specific developmental resource and that adaptive developmental responses arise due to the organismal regulation of development. The normativity of developmental processes can be accounted for by noting that norms emerge at the limits of mechanicism. This could be an organicist *dictum*. For instance, in autonomous systems, norms are defined with respect to the whole system and how each element plays a role in producing an adaptive outcome (contributing to self-maintenance) in different physiological processes. Norms cannot take place unless we do not posit that organismal activity is goal-directed toward the self-maintenance of the whole system. The limits of mechanicism refer to its incapability of explaining why something takes place because of the necessities of the whole system. Putting wholeness at the center of biology —as organicism does— means that taking the whole as a central explanatory unit is indispensable for tackling certain biological phenomena. I can extrapolate this to the developmental scale. Norms take place once we see developing systems pursuing a goal. The limits of mechanicism lie precisely in the fact that the self-regulation of biological information —a.k.a. interpretation— arises by treating the organism as the central explanatory unit of developmental processes. In this vein, the parts of a developmental system that participate in a developmental process must causally contribute to achieving the goal. As I will note in the following paragraphs, in the process of interpreting and using biological information, errors can arise and, as a result, a maladaptive outcome may be constructed.

It is central in any teleological theory to explain how errors are possible. If there are goals, there are norms, and therefore the possibility of error must be accounted for. As argued, error lies in the maladaptive regulation of developmental processes done by the agent. Certainly, a maladaptive response may arise, at least, for two reasons. The first reason concerns the fact that, in some cases, the biological information that a developmental system gets makes it difficult to produce an adaptive response. For instance, in some cases where genetic mutations arise, the organism may be incapable of accommodating such a change in order to produce an adaptive outcome. The same may happen if there are adverse environmental scenarios or other epigenetic changes. Surely, the most drastic, albeit quite common, case is when the system dies: the material conditions of the system do not allow to fulfil any of its goals.

The second reason, more interestingly for issues regarding regulatory processes, concerns maladaptive plasticity (Sultan, 2021). In these cases, the organism attempts to accommodate its developmental processes to the biological information available, but the developing agent nonetheless does not regulate adequately the

production of developmental outcomes. In general terms, we can recognize that maladptive plasticity can arise in two ways: in the interpretation of biological information and in the use of biological information. Note that these two options can be understood in a (informational) teleosemantic way: *misinterpretation* may arise when the agentive system does not properly integrate the pieces of biological information available coming from different sources, while *misuse* takes place when a proper interpretation is constructed but the developmental response is not adequate to the developmental conditions. In the teleosemantics' jargon, *misinterpretation* is an error in the production of a representation, in how the system represents the world —which is principally accounted for by producer-based theories, whereas *misuse* is an error in the use of the representation, in how the system uses the representation to produce an adaptive outcome —which is principally accounted for by consumer-based theories (cf. Section 3.3.2 for details).

In sum, the regulation of biological information does not always entail success, perhaps because the biological information does not enable the construction of an apt organism, or perhaps because the regulatory mechanisms fail at some point —in the intereptation or the use of biological information. The possibility of error entails the fact that maladaptive changes are possible even when goals are about adaptive change. This means that not all development outcomes are adaptive. This, surely, is good news for my proposal, insofar as teleological development must allow for the possibility of error. However, while not all developmental outcomes produced by the developing agent are adaptive, we may say that all of them are *adaptively directed*. The label *adptively directed response* captures the idea that even if a developmental outcome is the result of an agentive response, it may nonetheless be a maladaptive outcome. An adaptively directed response does not imply success. For instance, an organism with disadvantageous developmental resources may still be trying to produce an adaptive outcome, even if this is not possible. Goal-directedness, or the *effort* toward a goal —to use von Baer's expression— may still be present in those cases where no success is reached.

Adaptive developmental repertoires

Once normativity is introduced, we can represent repertoires in an adaptive way —i.e. not only as a repertoire of possibilities but as a repertoire in which each possibility has a specific adaptive value. This was a mandatory requirement stated previously. In the MS, adaptation is produced by populational forces moving populations within a design space according to their adaptive (trait fitness) values. The situation is similar at the physiological scale concerning the viability space. We must, therefore, see developmental repertoires as endowed with an adaptive assessment. In other words, different developmental outcomes in the developmental repertoire must have different adaptive consequences, in such a way that the

organism produces a specific outcome *for adaptive reasons*.

Figure 8.11 illustrates the adaptive character of developmental responses. The main difference between this picture and Figure 8.10 on page 268 is that the latter does not take into account the fact that developmental outcomes are regulated by a developing agent with its own goals and norms. Once normativity enters the scene, different outcomes within the developmental repertoire acquire different adaptive values. In Figure 8.11, I color-coded developmental outcomes to represent their degree of adaptiveness: Green means that the outcome is adaptive, yellow means that it is less adaptive, orange represents a maladaptive developmental outcome and red means death. Surely, this is a simple schematic picture, but it nonetheless captures the idea that developing organisms regulate their developmental trajectory according to the adaptive consequence of the outcomes produced.

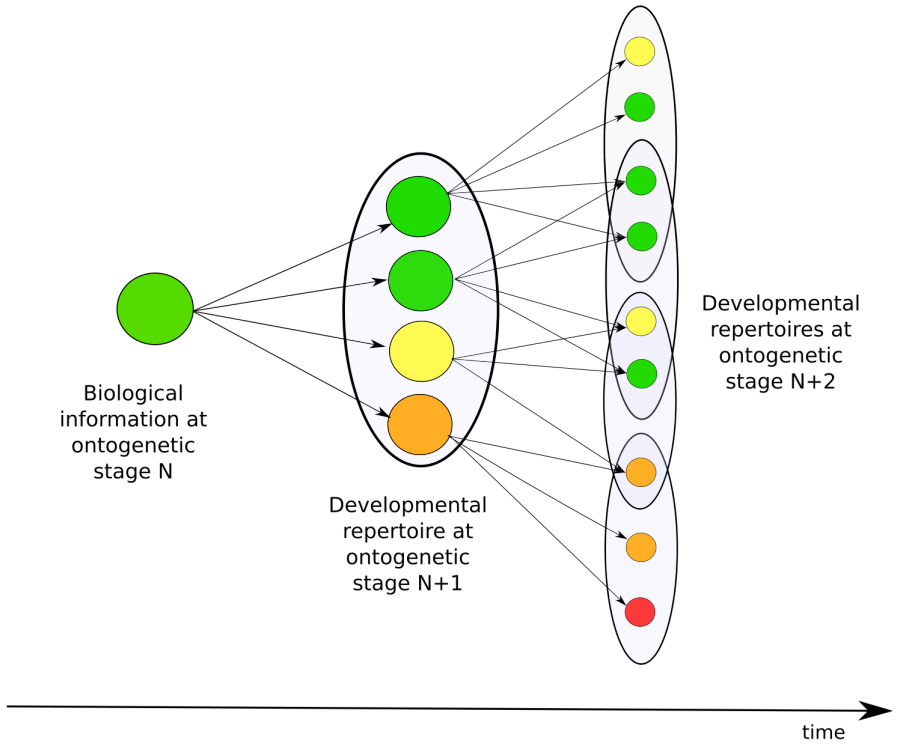

Figure 8.11: Adaptive developmental repertoires; see text for an explanation.

Moreover, in Figure 8.11, we can also highlight two other features. The first one concerns robustness. Robust developmental outcomes are those outcomes that can be obtained in different ways. In this sense, if two alternative developmental paths conduce to the same developmental outcome, such a trait is robust. Secondly, we can also appreciate the plasticity of developmental processes. From a

single configuration of biological information at ontogenetic stage N, the developing system has the potential to produce different traits according to the conditions of development (i.e. the available biological information).

8.2.4 Interim outline

Agential Teleosemantics defends that teleological explanations of development have the same causal structure as intentional explanations. This is the main reason for arguing that some sort of intentionality is present in development. How does Agential Teleosemantics conceptualize development as an intentional process? The main idea is that of an agentive system processing information to produce adaptive outcomes. This involves the core ingredients of an intentional explanation: using information to produce an adaptive (i.e. normatively evaluated) response. Once the first ingredient —information in developmental systems— and the second —agentivity and its relation with goals and norms— were introduced in Section 8.1 and Section 8.2 respectively, the core of Agential Teleosemantics is built.

All the elements of Walsh's Agential Perspective —goals, repertoires, agency, and experience— can be explained in informational terms. Moreover, I offered a number of remarks concerning the temporal dimension of teleological development. This allowed me to present all the elements of Walsh's perspective without assuming an adult-oriented stance. Developmental *goals* were defined as relative to ontogenetic stages and, concomitantly, the developmental *repertoire* was understood as the set of the possible developmental outcomes at a particular ontogenetic stage: the system *experiences* the U–I of the current ontogenetic stage. Finally, I supported a gradual and multi-level view of *agency*. Agentive properties emerge in development but are also relative to the level of analysis, in such a way that we can capture the fact that some form of agency is present at all developmental stages.

One of my motivations for presenting a teleological theory of development is that Walsh's theory does not provide an adequate analysis of invariance relations in development. My conclusion concerning his analysis was that invariance relations arise between a system having a goal and the means toward the goal. My purpose was to provide a theory that can understand such a relation in causal terms. As it was explained in Section 8.1.4, there are invariance relationships between biological information and developmental means. Having a goal is a property of agents endowed with biological information. Agential Teleosemantics explains how a system having a goal produces particular means toward the goal: an agent with a repertoire of possible outcomes at ontogenetic stage N uses its means to produce an adaptive developmental outcome at stage $N+1$. If the developing agent would have a different goal, then different means would be involved. The fact that having a goal precedes the means towards a goal is central to overcome the problem of

backward causation in teleological explanations: having a goal is a property of an ontogenetic stage N, and means concern the developmental changes towards ontogenetic stage $N+1$. Agential Teleosemantic takes advantage of the idea that information is the cause that imparts order in living organism, and attempts to connect this idea with the central dictum of the DT concerning the self-regulative capacities of agents.

8.3 The developmental path

In this last section I would like to discuss a number of relevant issues with the aim of making a stronger case for my analysis of teleological development in general and Agential Teleosemantics in particular. I will start in Section 8.3.1 by classifying the different sources of biological information in development in order to assess the explanatory role (and scope) of evolved information. Next, in Section 8.3.2 I will deal with some pending issues concerning different *desiderata* and challenges that Agential Teleosemantics must solve. Finally, in Section 8.3.3, I will suggest some historical reflections that may illuminate the place of Agential Teleosemantics within the DT.

8.3.1 Inheritance and the transmission of information

Is not Agential Teleosemantics excessively focused on development in order to explain the aptness of organisms? Should not evolution play any explanatory role? This subsection is motivated by these questions, and in order to address them I will refer to the work of James Griesemer on reproduction in the evolutionary process (Griesemer, 2000a, 2000b, 2014, 2016). I shall not offer a detailed introduction to Griesemer's framework here, partly because not all of his ideas fit in completely with my view, and partly because his analysis is not directly relevant to my presentation of Agential Teleosemantics. That said, I'm sure that further research may eventually result in a smoother integration of Griesemer's framework with a more streamlined view of Agential Teleosemantics. So, for the time being, I will essentially draw on Griesemer's work to present my own classification of developmental constraints. Griesemer is one of the main contemporary supporters of a developmental theory of inheritance and his theory of reproducers is a direct response to replicator views. In this context, he distinguishes between *inheritance* (the process of construction of traits) and *heredity* (the transmission of developmental resources). This distinction will be central in the setting up of an answer to the question that opens this subsection.

The classification of developmental constraints

I proposed to understand developmental constraints in terms of biological information. The idea of biological information fits in with the idea that developmental resources bring about a repertoire of discrete and finite variations. Griesemer's separation of heredity and inheritance, introduced above, corresponds to the difference between developmental resources and developmental processes. Developmental processes are about the use of developmental resources for the developing system to move on through the different ontogenetic stages.

In order to assess the role of evolutionary processes in this context, it may be useful to establish a classification of biological information into types. My conclusion would be that certain types of biological information have evolved and that this fact is explanatory central when addressing some important biological phenomena. I will also argue that not all developmental resources have evolved. To support these claims, it may be helpful to take a look at the different sources of biological information.

Following Griesemer (2000a), it is relevant to keep in mind the distinction between heredity and inheritance. Recall that the former concerns the transmission of developmental resources, while the latter concerns the process of development that produces cross-generational resemblance. The distinction between inherited resources and inheritance will help me provide a taxonomy of types of biological information that, despite its simplicity, nonetheless works for my purposes. To construct the taxonomy I will use my own labels —instead of adopting Griesemer's categorizations— to argue that the classification of developmental resources should be the one shown in Figure 8.12 on the facing page. Note that not all developmental resources are inherited. The first distinction is that some developmental constraints are constructed —they are the result of developmental processes— while others are transmitted —they are not brought about through developmental means. However, transmission here not only includes inheritance systems but other ways in which developmental resources reappear generation after generation.

The first class of developmental resources concerns any *biological information transmitted through systems of inheritance* (Section 4.2.1). For instance, genes are causal specifiers of development. DNA is a clear case where parents transmit developmental resources to offspring, but also epigenetic, cultural, and symbolic systems of inheritance supply development with biological information. In these cases, biological information operates as input to development, it is not produced by development. Obviously, it should not be forgotten that inheritance systems do not act only at the time of conception, but they provide causal specificity at different ontogenetic stages.

But inheritance systems are not the only reservoir of developmental resources. There are causal specifiers that are *not* part of the systems of inheritance and

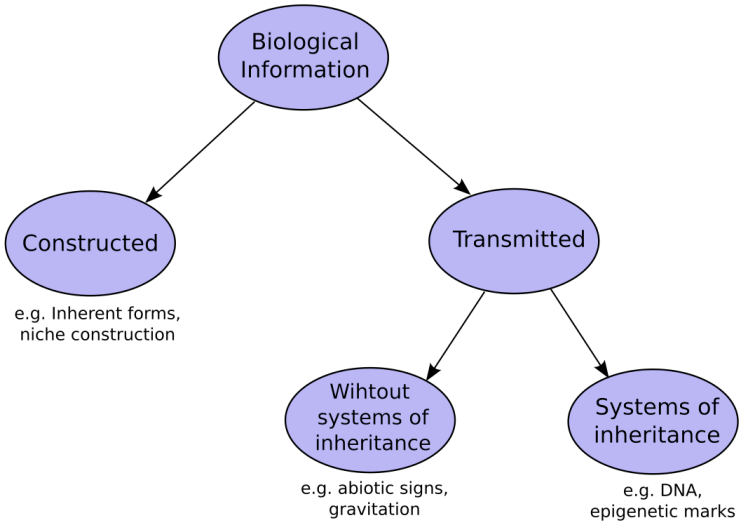

Figure 8.12: A general taxonomy of types of biological information.

that do *not* emerge during developmental processes. This class of developmental resources includes many things —mostly environmental, abiotic resources: abiotic signals, sounds, light, temperature, or gravitation, among others. These resources can also be understood in terms of biological information. One illustrative example of this in the recent history of DT is the study by Dubinin and Vaulina (1976, 47) on the impact of gravitation on morphology. Indeed, the idea that gravitational forces can be treated in terms of biological information is expressed quite explicitly in the paper: "*Variations* of gravitational field in any direction bring about numerous *changes* in organisms". Translated to our vocabulary of causal explanations, the authors assert that gravitational forces and organisms exhibit *invariance relations*: changes in the developmental resource (in this case, the directionality of the gravitational field) entail changes in the developmental outcome. Gravitation is, therefore, a causal specifier of development, it supplies biological information.

Finally, some resources are the result of ontogenetic processes and they are not transmitted. One example comes from inherency. As explained n Section 5.1.1, inherent forms are the result of the inner physical properties of developing systems; they have to do with emergent (morphological) patterns that arise through the dynamic interaction of parts of the system. The relevant point is that the parts of the system that interact to produce an emergent pattern arise during development. In this way, and aligned with the view presented in this section, the developmental outcome —in this case, a morphological emergent pattern — is the result of developmental constraints that were constructed during development. Ontogenetic niche construction offers another example. In this case, the develop-

ing system itself constructs its own resource from the environment and this may affect later developmental stages. An environmental developmental resource at a specific ontogenetic stage may be the result of a previous process of ontogenetic niche construction.

Surely, this classification deserves further elaboration and streamlining, but suffices to shed some light on the different posible sources of developmental constraints. Mine is a rather broad notion of developmental constraint, in consonance with my characterization according to which developmental constraints are those developmental resources that convey biological information. The definition of biological information in terms of causal specifiers, therefore, makes it possible to take into account different sources of developmental constraints. But this analysis is also helpful in the analysis of the relevant contributions coming from evolutionary history.

Evolution and biological information

I started this subsection asking whether my proposal is not too focused on development, neglecting the role of evolution in the explanation of adaptive phenotypes. I don't think so: my view on proximate causation as the *explanans* of aptness is motivated by the DT, the support to the Statisticalist School, and the importance of "taking development seriously" (Robert, 2004).

From the analysis of the previous section, we can also draw some conclusions, and quite expected ones at that, on the role of evolution in Agential Teleosemantics, which is *to provide 'rich' biological information —through inheritance systems— that has gone through previous selection processes.* As some causal specifiers (biological information) are transmitted by systems of inheritance, the idea is that biological information transmitted by systems of inheritance is 'rich' because such causal specifiers are the result of past selection processes.

First, note that one distinct character of developing systems is that their developmental resources have a high degree of specificity. Developmental constraints must make the construction of a phenotype possible. This possibility is enabled, in part, by the causal specifiers of development. This idea is also supported by Stotz and Griffiths (2017, 387), who argue that living systems "are structured so that many of their internal processes have an outstanding degree of causal specificity when compared to most non-living systems". My point is that such difference (between causal specifiers in living systems and causal specifiers in non-living systems) is the product of evolutionary processes. Causal specifiers in living systems are richer than causal specifiers in non-living systems, and such a difference is due, in part, to evolution. 'Richness' here refers to the degree of complexity in the possible outcomes within a repertoire. If a set of causal specifiers X enables a collection of complex organic forms and functions, X provides rich biological in-

formation. Conversely, if a set of causal specifiers Y enables a collection of simple forms and functions, then Y conveys simpler biological information than X. We can take for instance two animals with different degrees of organizational complexity: sponges and cows. Both have biological information that enables them to produce some developmental outcome. My point is that their different degree of complexity traces back (but not reduces) to their developmental resources —i.e. their biological information.

Evolution, therefore, is a process that results in the production of rich biological information. We already know how. Those forms and functions that maximize fitness increase the possibility of leaving more offspring. Leaving a higher number of descendants is not about transmitting traits, but about transmitting biological information through different systems of inheritance. Such biological information —among other developmental resources; cf. Figure 8.12 on page 279— makes possible (but does not determine) the resemblance of traits between parent and offspring. In this way, cumulative selection processes perpetuate through inheritance systems those causal specifiers that have contributed to produce complex and adaptive organisms. In doing so, biological information becomes richer.

The ontogeny of *intentional* information

The previous subsection may sound as an entirely mainstream and traditional way of understanding evolution: it sounds neo-Darwinian! The aim of this last subsection is to show that I am not contradicting myself supporting ideas that I have previously rejected. The same tension may arise if we look at the idea of the ontogeny of information, which states that developmental outcomes are not specified anywhere before developmental processes take place. In other words, that it is the very process of development that endows organisms with information. But if biological information is defined in terms of casual specifiers and, as I have just argued, many sorts of biological information are not constructed during ontogeny but transmitted through inheritance channels, it certainly looks like a contradiction. DNA is a paradigmatic causal specifier that conveys biological information and DNA is not obtained through ontogenetic processes but inherited at conception. The same applies to environmental and epigenetic modes of inheritance. In all these cases, the developing organism is not responsible for creating the information: Transmitted biological information *has no* ontogeny. So what is at stake here? Am I endorsing some kind of preformationist or replicator stance?

The answer to this question deserves a clarification concerning the thesis of the ontogeny of information. My proposal is the following one: the idea of the ontogeny of information should not be taken to be synonymous with the idea that biological information —understood in terms of G–S theory— cannot predate development. Biological information as causal specificity need not have an ontogeny. Rather,

I would like to put forth the idea that in the case of the *ontogeny of intentional information*, the determination of outcomes does not depend on biological/causal information, but on the *interpretation* of the whole informational network. Such process of interpretation is intentional insofar as it complies with the norms established by the network of biological information and results in the processing of such information towards the fulfilment of developmental goals. The ontogeny of intentional information presupposes that, even if certain pieces of biological information may be independent of developmental processes, the interpretation of such information is *always* a matter of ontogeny.

Biological information may not be intentional but its interpretation is. This also opens the way for an answer to the question about the normativity of signals. Since signals are also defined in terms of biological information, they exhibit no intentionality. Another way of looking at this conclusion is through a parallelism with emergent properties: emergent properties are ascribed to whole organisms but not to their parts. Similarly, as intentionality in development emerges at the level of the whole developing agent, it should'nt come as a surprise that its parts —biological information— have no intentionality.

Thus, biological information does not explain development. And the evolution of biological information is not about developmental processes at all. For instance, against Shea's Infotel theory, I have claimed that inherited biological information is not intentional information. Shea's etiological view defends that inheritance provides representations to be *read* by the developmental system in order to produce the evolved phenotype. The reasons why this (replicator) account is problematic were already exposed in Section 5.2.3. Shea's view contrasts with my view on interpretation, where *intentional* information is created during ontogeny, not during evolution. These labels —'to read' and 'to interpret'— illustrate very well the different perspectives. Reading a text leads to no space for creativity for the reader, but interpreting a text leaves room for different construals depending on the interpreter. If we translate this to the case of development, we see that, in Shea's view, reading does not require any kind of agency in developing systems: the text is already written by evolution and the developing system just reads it. But interpretation has to do with the fact that the result is not determined by evolved information and it (the result) crucially depends on the active participation of the agent that receives the information. Therefore, from the fact that evolution produces rich biological information we cannot conclude that this can explain developmental processes. This would be, as Quine (1976, 24) quipped, "to put the cart before the horse": to put replicators over organisms, biological information over interpretation, developmental resources over developmental processes, evolution over development.

That said, note that there is an idea in Shea's infotel theory that is indeed

supported by Agential Teleosemantics, namely that we can analyze the correlations between inherited information and phenotypic outcomes in informational terms. This is, to be sure, what biological information is about: measuring correlations in biological systems. The analysis of correlation does not give rise to any sort of dichotomies. We can certainly say that a trait strongly correlates with a particular DNA sequence, or that a trait exhibits variability in different environments, but these statements are not developmental explanations. Like Infotel, Agential Teleosemantics is informational teleosemantics (cf. Section 9.1.2). In opposition to Infotel, Agential Teleosemantics is not etiological —i.e. it does not define teleofunctions in terms of evolutionary processes. Evolution is central in the production of rich biological information which certainly enables, but does not determine, complex adaptive outcomes. The discovery of correlations between inherited information and phenotypic outcomes is clearly useful but it cannot be claimed to be an analysis of developmental processes.

8.3.2 Pending issues

The framework of Agential Teleosemantics presented so far has been built by borrowing and putting together a number of ideas about the nature of development. These ideas are not only about what development is but also about what development is not. Agential Teleosemantics has thus been formulated taking into account different conclusions that have been reached throughout this thesis: (i) that there are no $A - B - C - A$ chains in development, (ii) that a developmental goal cannot be preformed, (iii) that a developmental goal cannot be about adult states, and (iv) that a theory of developing organisms should depart from adult-oriented views. The unifying thread in this journey has always been the original idea of avoiding any resort to backward causation in teleological explanations. This has been achieved by a theory —Agential Teleosemantics— in which a system having a goal and its causal powers may be understood in naturalistic terms. As invariance relations involve a system having a goal and its means toward a goal, Agential Teleosemantic can construct teleological explanations without any sort of backward causation. That said, I would like to end this chapter by showing how Agential Teleosemantics is capable of abiding by the different constraints to teleology imposed during this thesis.

The Historical and Actuality *Desiderata*

In previous chapters, I introduced two important *desiderata*: the Actuality *Desideratum* (AD) (Section 5.2) and the Historical *Desideratum* (HD) (Section 3.2). The former states that intentional and teleological accounts must be capable of identifying the actual, intrinsic differences between a (token) intentional/teleological

system X and a (token) non-intentional/non-teleological system Y. The latter stresses that without a historical dimension concerning an adaptive bias operating in the past —be that in prior evolutionary or developmental stages— there is no teleological analysis in the present. I showed that autonomous systems theory accounts for the AD but not for the HD at the physiological scale, while the etiological theory accounts for the HD but not for the AD at the evolutionary scale.

Development seems to occupy an intermediate position, and the centrality of developmental processes that I conferred on the Agential Teleosemantic framework appears to make possible to comply with both *desiderata*. As for the HD, the history that defines the functions and norms of a system is its ontogenetic history. As explained in Section 8.2.1, as the developing system changes, its constitutive dimension gets modified. The generative process that produces a self-maintained system defines the normativity and the functionality of its parts. As for the AD, the norms and functions of a system are not based on a populational analysis but on the intrinsic capacities of the developmental system as a developing agent. While evolution is about the history of populations and physiology is about individuals, development is about the history of individuals.

Development and adaptive evolution

The DT has been promoting in the last decades a view of evolution where individual development plays a central role. The most categorical stance within this movement is epitomized by statisticalism: all causes of evolution are ontogenetic. The central elements of adaptive evolution by natural selection, inheritance, variation, and fitness, are caused by individual development (Nicholson, 2014; Walsh, 2007a); such causes are "The Proximate Foundation of Natural Selection" Stotz (2019, 343). In contradistinction with the MS externalist framework, here development is promoted to be the main responsible for creating new heritable variants that selection filters: "development is the artist, selection the curator" (Baedke & Gilbert, 2021). In this sense, if the causes of aptness are ontogenetic, we should be able to explain how the facts of development can be accounted for in teleological terms. Remember that teleological explanations are related to the explanations of aptness (Section 1.1.2). While other sciences such as physics or chemistry do not resort to any kind of functional, normative, or teleological talk, such categories appear to be what makes biology a distinct and unique scientific field. As noted, such categories are required to account for the aptness of living beings. Thus, if the causes of aptness are ontogenetic, and the aptness of living beings call for (naturalized) teleological explanations, then we must be able to explain in teleological terms how ontogenetic causes produce aptness in living beings exhibiting variation in heritable (individual) fitness.

In Agential Teleosemantics, the central ingredients of natural selection are integrated at the developmental scale. *Inheritance* is not divorced from development, but it is connected with how biological information is used by developing agents to construct traits with cross-generational resemblances. The *fitness* of individuals is the result of the adaptiveness of development. Individual fitness is obtained by the capacity of developing systems to regulate their material conditions of development. Agential Teleosemantics is about how such regulation could be adaptive. Finally, *variation*, one of the main worries of the DT, is an epigenetic phenomenon. The centrality of development in evolutionary theory makes even more important the distinction between the two ways of being of organisms. The *changing way of being* becomes particularly relevant in the quest for an ontogenetic theory of phenotypic variation. Even if a variant is caused by a genetic mutation, we already know that causal specifiers do not stand in a one-to-one correspondence with phenotypic outcomes, but that the connection between the genetic mutation and the phenotypic outcome is mediated by epigenesis and other causal specifiers. The main point is that whatever change occurs in developmental resources, the phenotypic outcome produced is always the result of integrating such a resource into the whole matrix of developmental interactions. Another way to express this is in informational terms: a variant in a causal resource produces new biological information, but the interpretation of this information (what actually connects the causal resource with the phenotypic outcome) is performed by the whole developing system during ontogeny.

Beyond the physiological level

At the end of Chapter 6, I presented three shortcomings of restricting our analysis to the physiological scale. I already dealt with all of them. First, I already explained how Agential Teleosemantics is capable of accounting for the Historical *Desideratum*, while this was not the case at the physiological scale. Second, an important neglected issue in autonomous systems theory is the problem of construction. How is autonomy reached? How are norms established? How do traits acquire different functional roles? The answer to all these questions is the same: through a process of construction. Goal-directed developmental processes construct the constitutive dimension of living systems and establish the norms that an autonomous system must obey in order to self-maintain. Agential Teleosemantics is the theory developed here to explain the construction of autonomy through ontogenetic processes.

The third shortcoming concerns the fact that the physiological scale has a too restricted explanatory scope; i.e. there are relevant phenomena in the explanation of aptness that cannot be treated at the physiological level. We have enough tools already to understand why this is the case: the double way of being of organisms

entails different goals, and therefore, different adaptive phenomena result from each goal. The goal of self-maintenance at the physiological scale does not produce the change in organic structure that we appreciate in development. This is why it is necessary to provide a theory of teleological development to account for those adaptive phenomena that are produced by developmental goals. The developmental scale of Agential Teleosemantics is necessary to account for a number of phenomena in biology connected with the emergence of adaptive changes in ontogeny, such as adaptive phenotypic variation, the stability of inheritance, and evolutionary novelties (Sultan et al., 2022).

Comparison with other proposals

Agential Teleosemantics is fully compatible with and motivated by the Agential Perspective defended by Walsh. This is made clear by the fact that the four elements of Walsh's proposal are also present in Agential Teleosemantics. First, as argued in Section 8.2.2, the notion of *agency* is central for understanding the normative dimension of informational processes in development. Second, the notion of *repertoire* is explicit in the gamut of outcomes that any causal specifier allows if we adopt a distributed stance on developmental resources. Third, the *experienced environment* is central in my explanation of goal-directedness in development by arguing that informational states represent the U–I of each organism. The main difference, as noted, lies in the last ingredient: goals. Agential Teleosemantics puts the emphasis on how *a system having a goal* causes certain means, an issue that remained obscured by Walsh's treatment of the relation between goals and means.

There are other proposals sharing some relevant elements with mine. For example, Godfrey-Smith (2017) also put forward a view of development based on the subjectivity of organisms and their capacity of experiencing an environment (c.f. also, Godfrey-Smith, 2016b, 2019). Following Lewontin's (1983b) insight, Godfrey-Smith recognizes in subjectivity a central causal factor in evolution, and such causal contribution is connected to the "proto-cognitive" capacities of "sensing and responding to events" (Godfrey-Smith, 2016b, 490), which lead organisms to adaptive responses according to the perceived environment. Godfrey-Smith thus unifies a number of central developmental phenomena, such as niche construction and plasticity, by invoking signaling systems as crucial ingredients in the proto-cognition of subjects: "Signalling, in its general causal pattern, can be seen as a combination of niche construction and phenotypic plasticity"; in other words, "a sender or producer of signs [an environmental cue] makes some change to the environment that is perceived by a receiver [the subject] who responds plastically" (Godfrey-Smith, 2017, 3).

Similarly, while the proposal of G–S was not directly linked to the issue of

teleology, Griffiths (2016) offered a sketch of how such a link could look like. Many ideas for Agential Teleosemantics were borrowed from this paper. His proposal is set up in contrast with Shea's evolutionary account of inherited representation and motivated by the search for an "ahistorical account of biological teleology" (Griffiths, 2016, 83). In this case, the ahistorical component just had to do with the absence of a phylogenetic history but it did not explicitly exclude an ontogenetic history: "With this ahistorical account of biological teleology, I can construct a definition of teleosemantic intentional information free of history. References to past evolution in Shea's conditions are simply replaced by references to *present evolution*" (Griffiths, 2016, 89; emphasis added). In his view, as is usual in any defense of teleofunctions at the individual scale, "the function of a biological trait is the contribution it currently makes to survival" (Griffiths, 2016, 83); thus Griffiths "identifies the functions of a trait with the features that are *adaptive*, whereas the historical [evolutionary] approach identifies functions with the features for which the trait is an *adaptation*" (Griffiths, 2016, 87; emphasis in the original).

Griffiths' emphasis on adaptive traits (at the individual level of analysis) in opposition to adaptations (at the populational level of analysis) leads him to argue for the idea of *adaptive information*. This idea is pretty much equivalent to the notion of normative information present in developing systems as defined by Agential Teleoseamantics. Like Agential Teleosemantics, Griffiths' proposal is also based on the individual causes that contribute to defining the fitness of an individual. Contrary to Shea's view that is blind in front of inner causal processes during ontogenesis, "the added value of the idea of adaptive information is that it can feature in proximate explanations of the operation of living systems" (Griffiths, 2016, 91).

8.3.3 A cognitive Revolution in Biology?

In this last section, I shall offer my musings on the idea of a Cognitive Revolution taking place in theoretical biology, not necessary because the DT promotes a revolution against the MS, but rather because the ideas defended by the DT are somehow similar to the foundational pillars of cognitive science raised during the Cognitive Revolution in the second half of the 20$^{\text{th}}$ century. This might help locating the position of Agential Teleosemantics within the DT, and also expose the inadequacy of the notion of affordance involved in Walsh's Agential Perspective.

Opening black-boxes

What does it means that theoretical biology is undergoing a Cognitive Revolution? This is not about arguing whether biology is transiting a revolutionary period or not. Rather, this claim is motivated by an analysis of the recent history of the Life

Sciences. My contention is that the current state of the art in theoretical biology is not too different from that of the disciplines affected by the Cognitive Revolution during the second half of the 20th century.[8]

The Cognitive Revolution gave rise to cognitive science as an interdisciplinary research field with the aim of understanding cognitive processes going well beyond the methodological and theoretical strictures of Behaviorism. Behaviorism is often characterized as having black-boxed the mind by just studying stimulus-response dynamics. Behavior was studied just by observing sensory inputs and biased behavioral outputs as if internal processes were explanatory secondary if not downright irrelevant. Somehow behavioral psychologists were able to see behavior through reinforced stimulus as if inner, cognitive processes were transparent.

The conviction that the black-box of the mind had to be opened sank in once science eventually recognized the importance of the internal processes responsible for connecting the input with the output (cf. Bermúdez, 2014, Ch. 1, for a textbook introduction to the emergence of cognitive science). Here, both mechanistic and intentional explanations have played crucial epistemic roles. Neuroscience contributed to a better understanding of the physiological mechanisms present in neural processes, while such disciplines as linguistics and cognitive psychology also changed their focus of attention from observable behaviors to internal processes. Since then (even though contemporary, post-cognitivist stances defy this view) cognitive science has been built on the idea that the internal processing of intentional information produces behavior. Getting under the skin opens new avenues for explaining the agentive capacity of cognitive systems and their related properties, such as rationality, decision-making, memory, and so on. So, the black-box was opened by elevating intentional information as the main causal factor in the production of behavior.

I see that a similar situation has been brewing in theoretical biology in relation to the status of the MS. I already explained why and how the MS black-boxed development. Supposedly, one can study adaptive evolution without looking at developing organisms. It is enough to look at genetic inputs and the adaptive bias operating on phenotypic outcomes. Evolutionary biologists could see phenotypic outcomes through genetic blueprints as if development were translucent (Walsh, 2000, 2007b). I already explained how the DT is casting the key that opens the box.

The crucial point, therefore, is that the same elements are present in both historical scenarios. There is a parallelism between Behaviorism and the MS in their central thesis: the black-boxing of internal processes, the claim that the explana-

[8] A similar contention is also found in Fodor and Piattelli-Palmarini (2010), Okasha (2018), Jablonka and Lamb (2020), and especially in several works by Ron Amundson (Amundson, 1988, 1989, 1990, 1994, 2006).

tion of aptness entirely relies on an input-output dynamics, and the assumption that the intrinsic and the agentive capacities of living and cognitive systems are both unexplained nor explanatory useful. Consequently, cognitive science opposes behaviorism in a similar way than the DT confronts the MS. So there is also a parallelism between cognitive science and the DT: both frameworks opened their boxes with the same key; i.e. by studying how the inner processing of information produces adaptive outcomes.

Having traced this historical connection, it is relevant to highlight that my conclusion is *not* that the DT promotes a revolution in biology. The issue of whether the MS deserves a full refurbishing, an extension or just a slight modification is controversial topic, as we already saw (cf. Chapter 4). Rather, my historical analysis is merely aimed at appreciating which ideas of the MS are challenged by the DT and, consequently, why the view promoted by an Agential Teleosemantics based on informational processing is relevant for the DT.

Representations and the limits of affordances

The parallelism between the Cognitive Revolution in cognitive science and the proposals of the DT in theoretical biology advanced in the last decades elucidates why Agential Teleosemantics deserves its own place within the DT. As explained, the core idea of the Cognitive Revolution is that internal processes matter; that what is going on inside the organism is relevant to explain its activity and, last but not least, that it is possible, from a scientific point of view, to understand such inner processes. DT also embraces this internalist spirit. The DT is about what is going on in organisms and about the evolutionary processes that organismal activities bring about. The central link between Agential Teleosemantics and the Cognitive Revolution is that both approaches put informational processes at the core of the explanations of organismal activities. Agential Teleosemantics, in sum, contributes some precious metal ore to the alloy for casting the key that opens the door to development.

Moreover, the link with cognitive science traced in this section also sheds some light on another issue. In Walsh's Agential Perspective, he appeals to the notion of affordance. As noted, such a notion comes from ecological psychology, a theory and research program within the so-called (radical) post-cognitivism. Post-cognitivism includes different approaches to cognition that challenge, in different ways, at least some central ideas of mainstream cognitive science. Ecological psychology, like another radical proposal, such as enactivism, rejects a cardinal idea in cognitive science: the very notion of representation. According to these projects, it is not necessary to posit (mental) representations in order to understand cognition. As a direct consequence of this stance, Walsh's view on teleology does not involve any sort of representational talk, insofar as it is hosted in and motivated by a radical

post-cognitivist notion: affordances.

What is at stake here is that the importance of internal informational processes is absent in Walsh's proposal and ecological psychology. The notion of affordance dispenses with any reference to how the organism processes the (synergistic) information to achieve a particular interpretation of its developmental conditions (de Llanza Varona, 2022). Accordingly, ecological information (the kind of information conveyed by an affordance) is directly perceived by the organism; no inner processing is needed. Alternatively, the representationalist view defended here, following de Llanza Varona (2022), explicitly assumes that distributed, synergistic biological information needs to be processed and integrated in order to achieve a representational state that conduces the organism towards an adaptive outcome. The notion of affordance (as it is understood in ecological psychology (Heras-Escribano, 2019)) may be appropriate to capture the fact that environmental information is always a matter of organism-environmental relationships: affordances are opportunities *for an organism* to act on its environment, as ecological psychologists put it. While this *dialectical* view of information is on the right track, in order to explain teleological development, affordances in developmental processes need be complemented with the different informational processes carried out by the developmental system; i.e. the notion of affordance does not explain central phenomena in development, and consequently, a representationalist view is needed.

My aim is thus not to argue against the notion of affordance but to argue that the need for informational processing in development requires anchoring the intentional character of teleological development in representationalist views of cognition, not in post-cognitivist ones.[9] This, consequently, is an advantage of Agential Teleosemantics: Agential Teleosemantics is motivated by representational views of cognition and the idea that information requires inner processing. In this way, in opposition to Agential Teleosemantics, Walsh's Agential Perspective, and the original idea of affordance, in particular, without further qualification do not seem to fit the bill for a full-fledged account of teleological development.

[9]This point requires some qualification. Orthodox cognitivism states that representations can be understood in symbolic terms in such a way that the mind can be explained as a sequence of (computational) operations over such symbols. This is not what I am defending here. Rather, I am referring to representational theories of cognition in a broad sense: the tenet that some notion of representation (i.e. semantic information, normative aboutness) is central to explaining cognitive processes properly.

8.4 Summary

The introduction of Agential Teleosemantics in this chapter was accomplished in two steps. First I argued that development proceeds by informational processing and defined the notion of biological information on the basis of the work of Paul Griffiths and Karola Stotz, which is in turn strongly influenced by the postgenomic view of development and Developmental Systems Theory. I also argued that the informational underpinnings of development are also present in other disciplines within the DT, such as eco-devo and biosemiotics. My conclusion was that biological information is a robust causal notion that captures the distribution of causal roles among different developmental resources in development.

The second step was to argue that informational processing during development involves some sort of intentionality. This conclusion follows from seeing development as a process guided by agents in order to fulfill their goals. I turned next to a definition of developmental goals in terms of the production of adaptive changes in the constitutive dimension of the organism at each ontogenetic stage. I then proposed to understand biological information in terms of developmental constraints in order to define the developmental repertoire as the range of possible developmental outcomes that may be obtained from a single configuration of biological information at a particular developmental stage. Developmental repertoires are established as a function of the path chosen by the developing agent. The presence of agency is not ubiquitous in development, but it is relative to the level of analysis, such that a full-fledged agentive system at the level of the whole developmental system only comes gradually into being during ontogeny.

I started Section 8.3 by classifying the different sources of biological information. This was crucial for understanding the importance of evolved information transmitted through inheritance channels, but also to put the emphasis on the true explanatory scope of evolutionary accounts. While evolved biological information is a central element in the explanation of aptness, it should not be confused with proper explanations of developmental processes. This lead me to clarify the notion of the ontogeny of information by stating that what emerges during development is not biological information itself —even though biological information can be constructed during ontogeny too, but developmental processes produce intentional information arising as a consequence of developing agents processing biological information; i.e. developing systems *interpreting the conditions of development*.

The motivation for presenting Agential Teleosemantics comes from the picture of development raised by the DT and different teleological accounts of development in the history of biology. To be sure, my proposal still needs to see its full adequacy to particular cases and to different kinds of organisms. For obvious reasons, this analysis falls beyond the scope of this thesis, but the acceptability of Agential Teleosemantics should be assessed first relative to its conceptual grounds,

over and above its eventual applicability to real developmental scenarios. With the aim of evaluating some conceptual issues related to teleological development, I therefore argued in Section 8.3.2 that Agential Teleosemantics is suitable for dealing with main challenges and *desiderata* put forth in this thesis, such as the Actuality Desideratum and the Historical Desideratum, the need to overcome the shortcomings of teleological physiology, and the importance of connecting teleological development with the causes of adaptive evolution.

What comes next is an application of Agential Teleosemantics to the cognitive domain in order to deal with Brentano's Problem. The analysis to be developed in the next chapter will hopefully also identify different classes of intentional systems. This will help, among other things, to pinpoint the kinds of intentional explanations present in developmental biology.

Chapter 9

Agential Teleosemantics and cognitive science

The goal of this last chapter is to extend Agential Teleosemantics to the cognitive domain. Such an extension suggests a new solution to Brentano's Problem, in addition to the need of extending the DT to cognitive science. So far, the focus of this thesis has been on the fields of the philosophy and the history of biology, not on the philosophy of cognitive science. The focus on the philosophy of biology is justified by the explanatory logic of teleosemantics, which encourages the adoption of the methodological principle of appealing to biological theory in order to solve some of the major problems in the foundations of cognitive science, like Brentano's Problem. A detailed analysis of Agential Teleosemantics in the context of Brentano's Problem partly requires connecting Agential Teleosemantics with the traditional debates turning around intentionality and the nature of representations that have been going on since the Cognitive Revolution, some of which appear enumerated in Section 1.2.3. This analysis, important as it is, falls however beyond the scope of this thesis. Thus, in this last chapter I will apply the explanatory logic of teleosemantics with the intention of extending the analysis developed in the previous chapter and offering a sketch of how Agential Teleosemantics would look like in the context of cognitive science.

The chapter is organized as follows. First, in Section 9.1, I will sketch the solution to Brentano's Problem that Agential Teleosemantics offers and assess the issue of content determination. In Section 9.2 I will start by arguing that both developmental explanations and cognitive explanations adopt informational teleosemantics. Once this parallelism is established, I will make a number of comments concerning the difference between intentional explanations of development, and intentional explanations in cognitive science. Finally, in Section 9.3 I will outline some ways in which the DT may be extended to the cognitive domain. Most biological approaches to cognition adopt a neo-Darwinian stance, but,

I shall argue, cognitive science may profit from the adoption of the conceptual and methodological tools developed in different areas within the DT.

9.1 Brentano's Problem in Agential Teleosemantics

In this section, I will use the conceptual apparatus of Agential Teleosemantics to present an alternative, non-etiological solution to the Problem of Brentano. As it should be clear by now, ontogeny will play a central role in my proposal for a solution. With this move, Agential Teleosemantics will come close to Dretske's theory. This is so for different reasons. First because Dretske's view, like mine, is based on ontogeny as the proper source of teleofunctions; and, second, because his is an informational teleosemantics, as we will be able to appreciate in detail presently.

What is the core of Agential Teleosemantic in cognitive science? My tentative answer to this question is this: development is the process of content determination and *the function of a representational system is to do whatever it has developed to do*. Before spelling out the details of this thesis, it is important to consider first two central issues for any teleosemantic project: the natural source of teleofunctions, and how representations acquire their content.

9.1.1 Goals and norms in cognitive systems

An important premise discussed in Chapter 7 concerns the double way of being of organisms. This led me to outline a view of teleological development where the main goal of a developmental system is that of producing adaptive changes at each ontogenetic stage. At each stage, the developing system changes its constitutive dimension while at the same time keeping its capacities for self-maintenance. This accounts for the fact that both ways of being are interrelated: teleological developmental processes construct the systems that teleological physiological processes must maintain at each stage of the life cycle. Following one of the tenets of teleosemantics —that cognitive functions can be understood in the context of a theory of biological functions, the view pictured in the previous chapter also promotes an analysis of cognitive functions, particularly in the case of teleosemantics, the functions of representational systems.

The difference between goals in living systems is also a difference between norms. On the one hand, the norms of a developmental cognitive system concern the adaptivity of the different changes that take place during development. On the other hand, as for *maintenance the way of being*, the norms of the cognitive system concern the role that each part plays in maintaining the organism in viable

conditions. As noted, a key idea in Agential Teleosemantics is that teleological development defines physiological goals and norms. In other words, by constructing the constitutiveness of an organism, developmental processes establish which norms a cognitive system must obey in order to maintain the system up and running.

As we saw in the previous chapter, this is central to my treatment of development as the source of cognitive teleofunctions. One of the shortcomings of autonomous systems theory concerning teleological physiology is that it remains silent about anything having to do with the construction and establishment of norms. Teleological physiology is about how a system, in this case a cognitive system, operates in order to self-maintain. This involves attributing norms to each part that participates in the operational relationships during physiological processes. However, it is necessary to explain where cognitive norms come from. Teleological development, unlike the evolutionary perspective defended by ET, is the proper source of cognitive teleofunctions. This is the core to solving Brentano's Problem in Agential Teleosemantics: a trait must do whatever it has developed to do.

Importantly, as it was explained in Chapter 8, the two ways of being may overlap. This is even clearer in cognitive development, which is a paradigmatic case of an open-ended process, where further cognitive capacities can be developed at different stages in the life cycle of the organism. Moreover, as in the case of physiological functions, representations play a crucial role in maintaining the adaptiveness of the system. In other words, like the heart is essential during embryonic development, representations are also important for the organism–environmental relationships during development, especially when further representational capacities will be acquired later on (as we will see in Section 9.1.2).

Recall that Agential Teleosemantics does not require a distinction between types and tokens as etiology does. This point is relevant because etiologists believe that only by distinguishing between types and tokens will we be able to explain norms in naturalist terms. Usually, this claim is a direct attack on Cummins-functions or Organizational Functions. According to etiologists, in both cases one of the problems is that we cannot explain where the normativity of the system comes from just by looking at what the organism is currently doing. In a nutshell, a historical dimension is needed, as acknowledged by the Historical *Desideratum*. Etiology argues that it comes from evolution, and therefore that a type-token distinction is needed. However, Agential Teleosemantic manages to provide an intrinsic theory of norms that does not require such distinction but that nonetheless rests on a historical analysis —ontogenetic history. So, instead of claiming that a trait token must do whatever its trait type determines (by evolutionary history), in Agential Teleosemantics, a trait token must do whatever it was determined to

do by its own ontogenetic history.

Swampman

There are many challenges to teleosemantics in the literature. However, a direct attack on the etiological theory of functions comes from the Swampman case. The intuitive idea is that Swampman should have norms and genuine representational content. This is a problem for etiology because etiologists claim that norms are determined by the evolutionary history: a *trait type coming into being during evolution*. Evidently, Swampman lacks this historical dimension. However, the proposal presented here may suggest a solution to this riddle. As explained, norms are established by the *individual process of coming into being*. What is distinctive of Swampman is that it is constructed abruptly, not by a gradual process. However, I think that Agential Teleosemantics can accommodate this case. The central point is that norms are established at the individual level, not at the populational one. Consequently, whatever process constructs an autonomous system, such a process would be responsible for establishing the norm. The fact that Swampman is not constructed by a process of gradual ontogenetic changes, does not necessarily entail that there is no process that determines its normativity.

We can appreciate better my idea if we consider not an imaginary case like Swampman but a real one like a physical dissipative system like, for instance, Bénard cells. Bénard cells are not autonomous systems: they lack the capacity of regulating the interaction with the environment and they have no distinct constraints playing different causal roles in the maintenance of the system. Consequently, Bénard cells have no genuine normativity. However, we can identify what are the necessary conditions to produce and maintain the system. These are not genuine norms —they are perhaps 'proto-norms'— in the sense that we can specify what must be the case in order to produce Bénard cells. The proto-norms of Bénard cells —or of any other physical dissipative system— arise spontaneously, like Swampman. Therefore, Bénard cells establish their proto-norms through the very process of their construction, even though the process is abrupt and not gradual. This analysis helps clarify my point. The difference between Bénard cells and Swampman is that the former have proto-norms, while the latter is assumed to have genuine normativity. However, Bénard cells are a real case where we can appreciate that their proto-norms are defined by the very process of construction, even though such process is spontaneous, as I alleged in the case of Swampman.

In fact, the implausibility of Swampman is related to the complexity of real ontogenetic processes. As noted, adaptive complexity requires an adaptive system that regulates *rich* —i.e. highly specific— biological information. So while it is not necessary to say that Swampman is physically impossible —insofar as it is one of the many 'thought experiments' in contemporary philosophy— I think that

Agential Teleosemantics can explain why a Swampman would have norms and genuine representational content, and at the same time we can explain why a real Swampman would be implausible: real dissipative systems that arise spontaneously are not autonomous agents, and real autonomous agents require a gradual process of construction where rich biological information is involved.

9.1.2 Content determination and biological information

I will focus now on another central issue for any teleosemantic project: how underived intentionality is naturalized; how representational content is determined without presupposing previous representational capacities. Here enters the scene a central key concept in Agential Teleosemantics: biological information.

At the beginning of Chapter 3, I introduced the referentialist-intensionalist debate for the explanation of content. I argued, on the basis of its naturalistic commitments, that teleosemantics is a referentialist theory. This means that the content of mental representations is not defined by some intrinsic relation between mental representations, but by environment-mind relations: content is determined by reference. This claim deserves some qualification, however. In Chapter 3 I also called attention to another important point: intensionalism may not be in complete opposition with referentialism. Naturalistic theories aim to explain underived intentionality —primitive representations. A referentialist theory is necessary for the project of naturalizing these representations. Therefore, the content of *underived* intentional states must be determined by their reference. However, when it comes to *derived* intentionality, perhaps intensionalist theories may turn out to be suitable and necessary. My aim now is to locate Agential Teleosemantics in the context of the referentialist-intensionalist debate.

At this point, it may be relevant to rehearse the reasons why underived intentionality is the main explanatory target in teleosemantics. Brentano's riddle is about how intentionality can be understood in naturalistic terms. This, in part, requires explaining how intentional capacities arise without presupposing previous intentional capacities. Underived intentionality, therefore, refers to those representational capacities that are explained without involving previous representations. If we naturalize underived intentionality, we would show that intentionality deserves its own legitimate place in natural science. This pinch of naturalized intentionality operates as the building block over which further representational capacities are constructed. Without this pinch, however, it seems that no naturalist project can be achieved. *Primitive representations* is a common label for those representations whose intentionality is not derived from other intentional stuff. Primitive representations thus work as the building blocks of representational development.

Primitive representations may involve different sorts of representations. A

common idea, even though not accepted by everyone (e.g., Fodor, 1998), is that primitive representations are *perceptual* —an expression that is synonymous with what others (e.g., Neander, 2017b, Ch. 2) call *non-conceptual*. This is connected to the referentialist underpinnings of teleosemantics, that is to the idea that the content of primitive representations is determined by their reference. If reference to the world is the main factor in the construction of representational capacities, then the obvious place to start is at the level of perceptual representations. Underived intentionality thus concerns the explanation of perceptual capacities without involving previous intentional capacities. While this does not guarantee our success in the setting up of a full-fledged naturalist project, naturalizing perceptual-primitive representations is nonetheless the first essential step.

My guiding idea is that the notion of *determination* in the explanation of content determination should be the same one as the notion of determination in Crick Information. Recall that determination in Crick Information concerns those developmental resources that exert fine-grained control in development; i.e. those resources that determine which outcomes may be constructed. Therefore, content determination rests on those developmental resources that determine which representations may be constructed.

Primitive representations

Why do perceptual representations in flies are different from perceptual representations in humans? As argued, the answer is ontogeny. Different ontogenetic processes produce different representational contents. The difference between two representational systems is, in part, a difference in the (neuro)anatomical structure of these systems. There is relevant biological information that participates in the development of the anatomical components of a representational system. Accordingly, the developmental resources involved in the development of the anatomy of the representational system are also causal specifiers in the process of content determination. This can be made clearer through an analysis of invariance relations. In general, changes in inheritance systems may lead to changes in representational capacities. Representations may change if some genetic mutation arises; in the absence of certain epigenetic marks different representations may be constructed; different cultures and social niches may elicit different representational capacities. To understand cognitive development is to see how these different developmental resources interact at different ontogenetic stages.

However, when the issue of content determination comes to the fore, a particular kind of biological information is primordial: *Environmental-Perceptual Biological Information* (EPBI). EPBI concerns those aspects of the environment that the organism interacts with during development. Among the different sources of environmental information, EPBI refers to a specific kind of organism-environment

interaction, namely the interaction through perceptual systems. The importance of EPBI is directly linked with the explanatory role of mental representations. Mental representations are about features of the world, in such a way that representations of the world elicit different behaviors that are apt according to what the representations inform about. So, although different kinds of biological information are involved in content determination (such as those involved in the construction of the anatomy of perceptual systems), EPBI concerns those elements of the world that are present during ontogeny and that primitive representations refer to. Insofar as the reference of a representation determines its content, the content of primitive representations is causally specified by the EPBI that participates in the acquisition of primitive representations.

EPBI, therefore, concerns those aspects of the world that interact through perceptual systems during the development of primitive representations. EPBI is present at all stages of cognitive development and it may be of different kinds depending on the perceptual systems involved. Different elements of the world produce inputs to the auditory system during development, for instance. The construction of the auditory system depends on such elements, in such a way that the perceptual representation constructed will refer to the elements of the environment that have participated in the development of the auditory system.

Primitive representations are thus acquired through organism-environment interactions. In particular, perceptual representations arise by interactions during ontogeny between the perceptual system and EPBI. This interactivism is about the different proximate causes involved in cognitive development, not about whether such causes are evolutionary or ontogenetic (as most dichotomic views assume). The normativity of primitive representations arises from the fact that primitive representations are constructed by a teleological process of cognitive development. Such a teleological process is about regulating the interaction of many developmental resources to produce adaptive changes during ontogeny. Ontogeny, as a teleological process, is responsible for creating normative content. The content of primitive representations is not preformed in the environment previous to developmental processes. Once the perceptual system —with its own anatomy and physiology— experiences its surroundings —a process that begins with the emergence of perceptual systems in pre-natal development (cf. Section 9.3.2)— the development of mental representations begins.

The ideas exposed so far are an outline of how content determination should work if we understand it as the result of ontogenetic processes involving different developmental resources. Thus, primitive representations have underived intentionality insofar as no mental representation operates as a developmental resource in their construction. In other words, what determines that a semantic item does not derive its intentionality from other semantic items is that none of the devel-

opmental resources involved are intentional.

Complex representations

In the previous subsection, I argued that the naturalization of underived intentionality comes from the development of representations bearing EPBI. Primitive representations are constructed on the basis of an ontogenetic history of environment-mind relationships based on biological information coming from the environment and perceived by sensory systems. Moreover, primitive representations deserve to be treated in intentional terms. They possess both aboutness and normativity. In the context of this subsection, such primitive representations fall under a referentialist theory: they are constructed via interactions between the representation and the reference throughout ontogeny. With this scenario set in advance, I propose to understand *the relation between referentialist-intensionalist approaches in terms of ontogenetic processes.* The construction of *complex representations* —those representations that are based on the internal relationship between contentful units— is an ontogenetic process that rests on previous primitive representations. Let me dwell on this idea a bit further.

A central motivation for adopting an intensionalist stance is that it seems implausible to argue that all representations can be acquired by environment-mind relationships. For instance, representations that refer to abstract entities such as DEMOCRACY, GOD or BITCOIN, seem not to bear any referential relationship with any particular item in the world. In general, it is not always the case that we need have a direct contact with the reference in order to acquire a representation. In other words, in many cases, EPBI about the reference of representations seems not to be always necessary.

I propose to accept both intensionalist and referentialist views of content but separated in ontogenetic terms. Further representations may be acquired not by environment-mind relationship but by the internal links between representations already acquired (and naturalized). This is the core idea of most neo-Fregean accounts. The main difference between neo-Fregean proposals and mine is that I attempt to provide a theory of underived intentionality from which further representational capacities grow. Note, importantly, that the construction of complex representational capacities involving internal semantic relationships is itself an ontogenetic process. As such, it can be captured by Agential Teleosemantics. The normativity of such representations is defined by an ontogenetic process. The reference of my representation of BITCOIN is determined by the (semantic) biological information that participates in the ontogenetic process that construct such representation.

The main difference between primitive and complex representations is the kind of biological information involved in their ontogenetic process. In particular, no

intentional information participates in the development of primitive representations —the most relevant biological information is EPBI. In the case of complex representations, however, the primitive representations operate as biological information for further development. As primitive representations participate in the development of further and complex representational capacities, then *primitive representations also convey biological information to development*. Consequently, while other sorts of biological information are relevant in the complex representation, for my proposal, however, it is enough to say that in the development of complex representation *some* biological information comes from representational capacities previously acquired; i.e. that complex representations rest on intentional biological information (primitive representations).

The different sources of biological information in the determination of representational content are illustrated in Figure 9.1 on the following page, in the context of the classification of biological information presented in Figure 8.12 on page 279. Figure 9.1 shows that primitive representations are constructed biological information based on other biological information (EPBI and other developmental resources). Therefore, no semantic capacity is involved in the development of complex representations. Moreover, complex representations rest on different sources of biological information, and among them, the most relevant one is that of primitive representations; consequently, a constructed semantic capacity plays a role in the development of complex representations.

9.2 Informational Teleosemantics in Life Science

I stated that teleological explanations of development involve some sort of intentionality. My aim now is to clarify the connection between intentional explanations of development and intentional explanations in cognitive science. I will start by noting a point in common, namely that both kinds of explanations are informational. *Informational teleosemantics* is involved both in the explanation of developmental processes and of cognitive processes: the idea that informational processes conduce systems towards the achievement of goals. In this sense, information and teleology are common grounds in explanations of development and cognitive systems. Next, in Section 9.2.2, I shall offer some reflections on the possible difference between intentional explanations in cognitive science and development. My aim is not to provide any definitive answer to this (hard) problem but mainly to suggest some reasons why physiological categories and mental representations should *not* be extended beyond the living realm.

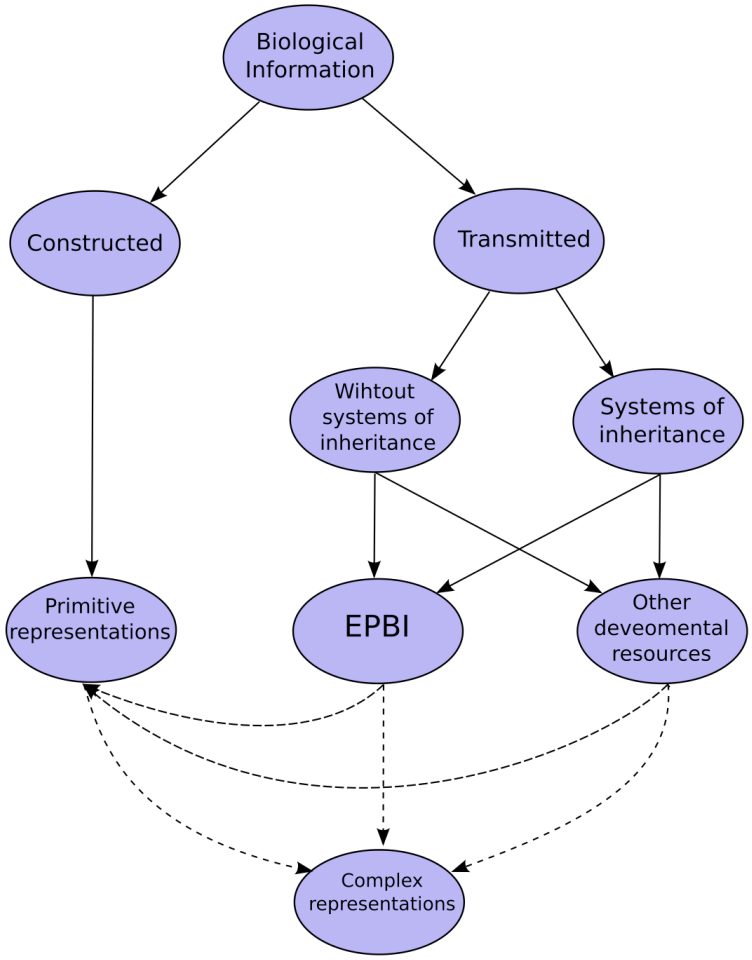

Figure 9.1: A general taxonomy of biological information involved in content determination. Dotted lines between two items indicate that one participates in the construction of the other.

9.2.1 Agential Informational Teleosemantics

Agential Teleosemantics is informational. It is based on the process that produces a certain outcome and on the fact that such a process is informational. In development, biological information is processed to produce an outcome; such an outcome deserves a teleosemantic treatment insofar as intentional information is involved. Therefore, Agential Teleosemantics, by emphasizing its informational underpinnings, may operate as a link between biological development and cognitive science. In other words, Agential Teleosemantics may contribute to the unification of informational talk in the Life Sciences. There is already a long tradition of informational teleosemantics, starting with the works by Stampe (1977) and Dretske (1981, 1988) and the more recent ones by Shea (2007a), Martínez (2013b), Neander (2017b) and Artiga (2020), among others. I will first present the core idea of informational teleosemantics and next I will argue that Agential Teleosemantics in development is also informational. This will help reinforce the link between the disciplines of cognitive science and developmental biology and the claim that unification of informational talk in one and the other is possible.

Informational Teleosemantics in cognitive science

Neander's (2017b) book is one of the most robust and comprehensive defenses of informational teleosemantics. Neander supported an etiological account, but one of the virtues of her book is that she presented an argument in favor of informational teleosemantics that is independent of any biological theory of function. In this way, her argument opens the door for a different, non-etiological teleosemantics. Her argument, which she called *The Methodological Argument for Informational Teleosemantics* (Neander, 2017b, 73-74), runs as follows:

P1 A notion of normal-proper function is central to the multilevel componential analyses (i.e., functional analyses) of the operation of bodies and brains that are currently provided by physiologists and neurophysiologists.

P2 The brain's normal-proper functions include cognitive functions.

P3 The same notion of function (mentioned in P1) is central to the functional analyses of cognition that cognitive scientists provide.

P4 An assumption in the mainstream branches of cognitive science is that cognition involves information processing.

P5 The (relevant) notion of information involved in such talk of information processing in cognitive science is (not a semantically evaluable notion but instead) a notion of natural-factive information.

P6 Cognitive science posits 'normative aboutness', with the norms derived from the aforementioned normal-proper functions and the aboutness from the aforementioned natural-factive information.

C Informational teleosemantics is supported by the explanations of cognition that the mind and brain sciences currently provide.

The first three premises (P1, P2, and P3) state that some of the functions of the brain are cognitive normal-proper functions. In other words, that cognitive functions are one kind of the functions that brains perform, and such functions involve normative talk. The methodological character of Neander's argument is based on observations of what cognitive scientists do; i.e. on how contemporary cognitive science goes about explaining and thinking about cognition: the first three premises are mere descriptions of actual scientific practice. The fourth premise (P4) states that the mainstream paradigm within cognitive science is based on the idea that cognition is about information processing.[1] Crucially, informational talk in cognitive science is primarily concerned with the causal, factive notion of information (P5). The quantum leap towards informational teleosemantics consists in joining together the first three premises with the fourth and the fifth: the first three state that *normativity* is involved in cognitive science, and the fourth and the fifth state that the notion of *aboutness* comes from information processing. As a result, premise six (P6) claims that cognitive science involves *normative aboutness*. The conclusion (C) that follows is that informational teleosemantics stands (or falls) to the extent that it can explain both aboutness (through information processing) and normativity (through biological function).

The differences observed in different teleosemantic projects may have different sources, but, if we pay only attention to their biological groundings, we shall see that different notions of biological function and natural normativity lead to distinct teleosemantic projects —i.e. different theories of function explain the first three premises differently, and Agential Teleosemantics has its own particular way of explaining normal-proper functions in cognitive science.

Informational Teleosemantics in developmental biology

We can recast Neander's argument in developmental terms. It takes the following form:

[1] Note however that not everyone involved in the cognitive enterprise adopts an informational view of cognitive processes. Different approaches within (radical) post-cognitivism, such as enactivism, would reject informational talk. In this sense, Neander's argument justifies the need for informational teleosemantics in mainstream cognitive science, but not necessarily in some alternative and emergent views of cognition.

P1 A notion of normal-proper function is central to the multilevel componential analyses (i.e. functional analyses) of the operation of developing systems that are currently provided by different areas of the DT.

P2 A developing system's normal-proper functions include adaptive and regulative functions.

P3 The same notion of function (mentioned in P1) is central to the functional analyses of development that the DT provides.

P4 An assumption in some branches of the DT is that development involves information processing.

P5 The (relevant) notion of information involved in such talk of information processing in development is (not a semantically evaluable notion but instead) a notion of natural-factive information.

P6 DT posits 'normative aboutness', with the norms derived from the aforementioned normal-proper functions and the aboutness from the aforementioned natural-factive information.

C Informational teleosemantics is supported by the explanations of development that the DT currently provides.

The first three premises establish that some sort of normative talk is involved in the regulative and adaptive capacities of developing systems. As explained in Chapter 8, such normativity comes from the goal-directedness of developing agents. Moreover, like Neander's, my argument is also methodological: it takes insight from how the DT explains development in adaptive terms. In this context, I mentioned, at least, three research areas that incorporate the notion of information: G–S's proposal motivated by post-genomics and the DST, eco-devo, and biosemiotics. These research areas operate as the backup of premise four. The notion at play, as the fifth premise states, is the causal notion of information: what was defined as biological information on the basis of the framework developed by G–S. Informational Teleosemantics in development arises by merging the agential capacities of living beings with the informational talk in developmental biology. As a result, the DT posits –albeit implicitly— (some sort) of normative aboutness. To explain development in adaptive terms, the DT uses informational teleosemantics.

I propose, therefore, that Agential Teleosemantics may be the appropriate tool for unifying different kinds of informational talk in the Life Sciences, particularly, in cognitive science and developmental biology: both areas involve a causal notion of information and normative assessments. My argument in this section is about

the *similarities*, claiming that both domains involve normative aboutness, but, what about the differences between intentional explanations in development and cognitive science?

9.2.2 All the way down

Should cognition go "all the way down" (Levin & Dennett, 2020)? The contention that all organisms have cognitive or mental capacities is certainly a controversial one that has given rise to a heated debate within the contemporary philosophy of biology. I would like to briefly comment on this debate, not with the purpose of taking sides, but just to elucidate a few points of my proposal concerning intentional explanations of development and cognitive science.

The thesis that any living being has cognitive capacities (the Life-Cognition Continuity Thesis) is defended by different contemporary thinkers (e.g., Bechtel & Bich, 2021; Godfrey-Smith, 2016a; Kováč, 2006; Levin, Keijzer, Lyon, & Arendt, 2021; Lyon, 2015; Maturana & Varela, 1980; Van Duijn, Keijzer, & Franken, 2006, among others). A similar, yet not necessarily equivalent, thesis is the Life-Mind Continuity Thesis: all living beings have a mind and all systems with a mind are alive (Dennett, 2017; Godfrey-Smith, 2016b; Thompson, 2007; Wheeler, 2011; Wiese & Friston, 2021). The Life-Cognition Continuity Thesis is usually taken to be equivalent to the Life-Mind Continuity Thesis; however, one could argue that some cognitive processes (let's say, perception) are not mental processes. Another closely related thesis is the Life-Agency Continuity Thesis: all living beings are agents and all agents are living beings. As before, the equivalence with the previous theses is in this case relative to the co-extensionality of mind-cognition-agency.

In the wake of autonomous systems theory and Walsh's agential perspective, for obvious reasons, here I supported the Life-Agency Continuity Thesis. In this context then, the following question immediately arises: Is Agential Teleosemantics also committed to the Life-Cognition Continuity Thesis? Admittedly, many phenomena in which some kind of agency appears to be involved suggest the presence of a sort of cognitive or 'proto-cognitive' properties, regarding for example the sensorimotor capacities of developing organisms and the regulative character of developmental and physiological responses. I shall not discuss the extensive literature on this topic referred to above because a detailed analysis of this issue falls out of the scope of this thesis. Let me nonetheless point out that an answer to the previous question involves resolving a number of terminological issues. To argue for the co-extensionality of agency and cognition depends on how these terms are interpreted and the criteria used to interpret them, such as biological criteria (evolutionary, developmental, functional, morphological) or philosophical criteria (through epistemic analysis of biological explanations or reflections on natural kinds in biology).

But what about the commitment of Agential Teleosemantics with the Life-Mind Continuity Thesis? Clearly, no strong (or weak, for that matter) commitment to any sort of Life-Mind Continuity Thesis is necessary for Agential Teleosemantics to work. The principal reason concerns the fact that the kinds of representations involved in cognitive science are typically labeled *mental* representations, but it is not necessary to argue that the kinds of representations involved in developmental explanations are mental. Agential Teleosemantic is just committed to arguing that some notion of representation arises in development as the result of agents using biological information for adaptive reasons. In the context of Agential Teleosemantics, the Life-Mind Continuity Thesis is also a question about the nature of representational systems present in developmental and cognitive processes. As noted previously, I am not going to get into this difficult (and old) issue, nor will I provide a taxonomy of representational systems (as, for instance, Dretske, 1988 did), I just intend to argue that the analysis of different sorts of representational systems in nature may help evaluate the Life-Mind Continuity Thesis. In this sense, the Life-Mind Continuity Thesis can be approached, for instance, by asking whether paradigmatic properties of mental representations (e.g. generalization, abstraction, systematicity, compositionality) are present or not in teleological explanations of developmental processes. Similarly, we can ask about the properties of informational systems in developmental and cognitive systems. While I argued that both developmental and cognitive processes can be understood in an informational way, there might be relevant differences between them that support the idea that mental representations are exclusively present in cognitive processes. While I will not address this issue here, there are two reasons to believe that representational systems in development differ from representational systems in cognitive science.

First, as noted, the content of the Life-Mind Continuity Thesis clearly hinges on how 'mind' and 'life' are understood. If 'mind' is understood as necessarily involving psychological categories, then it is clearly the case that Agential Teleosemantics can get through without positing any sort of psychological powers in developing agents. While it is not clear whether a term is used in a metaphorical sense or if it refers to a real natural kind, Agential Teleosemantics seems to dispense with form psychological labels in order to understand developmental processes —except in the case of cognitive development, surely. Second, the issue of the metaphorical or heuristic value of some expressions was already addressed when discussing the 'shortcut strategy' in Section 7.3.2. In that section, I revived an key insight by von Baer concerning the importance of the terms used in teleological explanations. My analysis concluded that Agential Teleosemantics should not naturalize any sort of purposive locutions —paradigmatic of psychological states— and just deal with the relation between a system having a goal and the means towards a goal: Agen-

tial Teleosemantic needs not resort to any kind of psychological locution to account for the connection between having a goal and a means.

9.3 Extending the developmental turn

While discussions in this thesis primarily turned around the field of philosophy of biology, there are many important debates in cognitive science related to intentionality that were not addressed here. However, my purpose here is a different one. The aim of this last section is mostly programmatic. I will discuss some ways in which debates within the philosophy of biology may have impact on the cognitive sciences. The motivation for this is that most approaches within the cognitive sciences assume a neo-Darwinian framework, therefore the emergence of an alternative consensus in biological theory forces a reconsideration of all these approaches in the light of this new consensus: new approaches to understanding cognition are possible.

Indeed, when cognitive scientists adopt a biological framework to theorize about cognition, neo-Darwinism appears to be the only game in town. Neo-Darwinian traces are present mostly in theories of cognitive development. In this context, dichotomic thinking is ubiquitous, as is the use of genetic metaphors and the idea of evolutionary design. Table 9.1 on the facing page is a compilation of some relevant references in cognitive sciences where neo-Darwinism is (sometimes implicitly) present.

The conceptual and methodological tools of the DT however have so far had little impact on the cognitive sciences, although some proposals incorporating a number of insights from the DT (collated on Table 9.2 on page 310) have already been put forth. This section is aimed at identifying some ideas of the DT that may prove relevant for the study of mental representations, and which are also connected with Agential Teleosemantics.

9.3.1 Where representations do not come from: against (old) semantic preformationism

On Table 9.1 on the facing page, we find a number of references where dichotomic thinking is present in some way or another. The fundamental question in most of these works is: where do representations come from? Some argue that they come from evolution, others from the experience, and in general, a consensus with some form of interactivism is generally observed: representations come from the interaction of evolved capacities with the environmental context of development. While I already criticized this view in Section 5.2, it is relevant to note that the

Field	References
Teleosemantics	Dretske (1988); Millikan (1984); Neander (2017b); Shea (2018)
Philosophy	Carruthers and Laurence (2005, 2007a, 2007b); Cowie (1999); Fodor (1981, 1998); Margolis (1998); Margolis and Laurence (2011); Plotkin (1994); Prinz (2002)
Psychology	Carey (2009); Dehaene (1997); R. Gelman and Carey (1991); S. A. Gelman (2003); Marcus (2004); Spelke (1998, 2017)
Linguistics	Berwick and Chomsky (2016); Chomsky (2000); Tomasello (1999)
Evolutionary Psychology	Barkow et al. (1992); Mercier and Sperber (2017); Pinker (1994, 1997)
Sociobiology	Buss (2008); Wilson (1975)

Table 9.1: Neo-Darwinism in Cognitive Science: a (non-exhaustive) list of fields and literature in the cognitive sciences committed with or motivated by the neo-Darwinian framework.

issue of the origin of representation surely predates the very notion of representation; in other words, dichotomic categories in cognitive science were inherited from old disputes between nativism and empiricism in the history of philosophy (Cowie, 1999; Lorenzo & Longa, 2018; Plotkin, 2008; Rama, 2018).

The important point is that the argument present throughout the history of philosophy is quite similar to the argument present in cognitive science. The empiricist tradition, since Aristotle, the British Empiricists, and the Logical Positivists, among others, bequeathed the conceptual apparatus underpinning contemporary empiricism. In particular, the core idea, labeled the Perceptual-Priority Hypothesis by Prinz (2002, 106), is the one encapsulated in Aquinas' famous statement: "Nothing is in the intellect that is not first in the senses". Representations come from the environment, while the mind is a *tabula rasa*, it is sensory inputs that bestow cognition with representational capacities. As for the nativist camp, Chomsky's famous *Poverty of Stimulus Argument* to posit innate linguistic capacities traces back to Plato's Reminiscence Theory. The idea that "the exterior is not enough" to explain the presence of cognitive capacities was understood as an argument in favor of nativism. This explanatory logic is also present in the Core Knowledge Hypothesis in psychology, assumed for instance by Spelke and Kinzler (2007) and Carey (2009). In like manner, we can find a similar version of Leibniz's *Impossibility Argument* in Fodor's *Mad Dog Nativism* (Fodor, 1998; Rey, 2014),

Field	References
Teleosemantics	Bickhard (2003, 2009a)
Philosophy	Bateson and Mameli (2007); Linquist (2018); Lorenzo and Longa (2018)
Psychobiology	Gottlieb (1998, 2002); Michel (2013); Michel and Moore (1995); C. L. Moore (2003)
Linguistics	Balari and Lorenzo (2013, 2015, 2018); Balari, Lorenzo, and Sultan (2020)
Evolutionary Psychology	Caporael (2001); Stotz (2014)
Niche Construction	Bertolotti and Magnani (2017); Heras-Escribano and de Pinedo-García (2018); Kendal (2011); Kerr (2007); Sterelny (2010); Stotz (2010)
Evo-Devo of Cognition	Amundson (2006); Balari and Lorenzo (2021); Finlay (2007); Ploeger and Galis (2011)

Table 9.2: The Developmental Turn in Cognitive Science: a (non-exhaustive) list of fields and references in the cognitive sciences committed with or motivated by the DT.

which states that innate representations are those that are theoretically impossible to be learned.[2]

Agential Teleosemantic departs from these disputes. It is not about finding a mid-term position between nativism and empiricism, in such a way that the positive points of each side of the debate are captured. This would be just a standard interactivst view —or a Type 1 interactivism, as presented in Section 5.2— which has "not been enough to drive away the ghost of dichotomous views of development" (Gray, 1992, 172). Agential Teleosemantics states that the problem is not the answer but the very question about where representations come from. This question hides a sort of semantic preformationism that is then present in the respective answers —empiricism and nativism. The idea that representations come from somewhere, that they exist independently of developmental processes is the core of a preformationist attitude towards (semantic) information: "it exists before its utilization or expression" (Oyama, 2000b, 2).

The alternative to the old semantic preformationism must be some form of constructivism: how representational capacities are constructed throughout developmental stages? This question looks for the proximate causes of cognitive development assuming that semantic information does not predate the process of construction but that it results from ontogeny itself; this is in a nutshell the idea

[2]See Cowie (1999) for an analysis of the contemporary versions of Plato's Poverty of the Stimulus Argument and Leibniz's Impossibility Argument.

of the ontogeny of semantic information. Note that, as I shall point out later, this constructivist view of representations is indeed the same one defended by a developmental conception of inheritance. In this vein, insofar as "the most important tenet of the constructionist view" is that "all phenotypes are constructed, not transmitted" (Gray, 1992, 177), I conclude that all representations are constructed, not transmitted by inner or outer channels of inheritance.

This constructivist view of representations has a direct connection with two central areas within the DT: niche construction and eco-devo. Regarding the former, recall that there are different conceptions of niche construction. It is important to distinguish between the individual and the populational conceptions of niche construction. Within the individual conception, there are two sorts of phenomena: first, *material niche construction* —concerning external environmental changes brought about by the organism— and, second, *experiential niche construction* —regarding the way in which each organism constructs its own experience of the environment. *Construction*, in the experiential reading, is an epistemic notion. This epistemic reading links niche construction to the issue of the acquisition of representations: explaining how a cognitive system epistemically constructs its niche is the main aim of a constructivist view of representation that departs from the preformationist legacy. In the wake of this, niche construction should also be extended to the realm of cognition, as some scholars have already advanced (Bertolotti & Magnani, 2017; Heras-Escribano & de Pinedo-García, 2018; Kendal, 2011; Kerr, 2007; Sterelny, 2010; Stotz, 2010).

Now, the issue of the development of representations can also be seen as part of eco-devo. Eco-devo is primarily concerned with organism-environment relations throughout ontogeny. In this context, Sultan (2015) has introduced the notion of *environmental-cue/system-response* to emphasize the role of exogenous inputs in the organismal regulation of development. In this vein, an eco-devo perspective of cognition may help in the search for the exogenous causes of cognitive development and how these interact with the inner dynamics of development. This would imply extending the environmental-cue/system-response dynamics across the whole living domain, and assuming that "[i]n a broader framework, all types of environmentally mediated phenotypic expression [representational capacities included] can be viewed as cue and response systems" (Sultan, 2015, 51). I already argued that primitive representations are generated through the interaction between perceptual systems and EPBI during development. In this framework, we can integrate representational development into Sultan's proposal by considering EPBI as information coming from the environment and cognitive systems adaptively responding to it.

9.3.2 Prenatal development and the myth of birth

Birth seems to be the landmark of all innate vs. acquired assessments. Surely, different diagnostic properties for nativism are typically invoked, as we saw in Section 5.2. But one of the paradigmatic properties of an innate trait is that it is present at birth. This fact becomes particularly relevant when we consider the development of mental representations. Mental representations play a crucial role in endowing a cognitive system with the capacity for adaptive interactions with the environment. Mental representations allow a cognitive system to experience an environment and adaptively respond to it. By taking birth as the 'birth of experience', we obtain a criterion for innateness: if a system has a representation before the system has had experience with the reference of the representation, then this representation must be innate —it does not come from the experienced environment. If a system represents certain features of the environment before the system has experienced such features, then such representation must come from somewhere else, but certainly not from the environment, not from experience.

There is a long tradition of psychological studies of prenatal development, culminating in contemporary developmental psychobiology (Michel, 2013; Michel & Moore, 1995), that has always fought against the view described in the previous paragraph.[3] The principal break is with the very idea that birth entails the birth of experience. Studies in prenatal development evidence that much is taking place before birth. In this sense, what seems to be an innate representation must have had a history of ontogenetic processes.

The importance of prenatal development is connected with the Phylogeny Fallacy and the explanatory vacuity of dichotomies. As noted, the main problem is that saying that a certain trait is innate does not inform us about its development. It says nothing about how such a trait came into being. Consequently, employing dichotomies "serves to block the investigation of their origin just at the point where it should leap forwards in meaningfulness" (Lehrman, 1953, 346). This is Kuo's lesson that we discussed in Section 5.1.1, namely the fact that no real developmental explanation is given in a dichotomic frame, and that "to call an acquired trend of action an instinct is simply to confess our ignorance of the history of its development" (Kuo, 1921, 650), but, if we acknowledge the relevance of prenatal cognitive development, such a history of development might be revealed.

The result is that birth as the beginning of experience is a myth: "a person's birthday is really just another day in the continuous series of days that stretch from conception to death" (D. S. Moore, 2001, 186). What is present at birth must come from somewhere else, since "pointing out that a pattern of behavior is

[3]The field of (cognitive) neuroconstructivism, led by Karmiloff-Smith (2006, 2009), also departs from the traditional categories of development in its quest for the causes and processes of cognitive development.

present at birth does not explain how it comes to be present at birth" (Johnston, 1987, 178–179). But we cannot explain whatever is present at birth in evolutionary terms, or to posit genes for those traits. This would not explain the actual process of *coming into being*.

9.3.3 Further directions

In this last subsection, I shall consider three other research areas within the DT that may nurture the study of cognitive development. My aim is not to provide a well-defined agenda, but just to identify a number of conceptual and methodological tools that could contribute to a change of perspective and stop looking at cognition with (only) neo-Darwinian glasses in the quest for the causes of cognitive development.

Inheritance

The developmental conception of inheritance has contributed some relevant theoretical and empirical insights. On the theoretical side, remember first that inheritance is seen as the process of trait construction. Inheritance is responsible for explaining resemblance. Consequently, as content is determined by the process of trait construction, inheritance is also responsible for explaining the resemblance of representations in different individuals. In a nutshell, if content is determined by developmental processes, two representations would have the same content if they share the same developmental trajectory (cf. below for some ideas concerning sameness). The idea of resemblance, connected with a developmental conception of inheritance, may enrich different perspectives concerning the objectivity of representational content. Ever since Descartes the mind has been the territory of subjectivity, but the rise of cognitive science opened the door to a scientific inquiry of mental processes. In other words, the possibility of explaining the content of mental states in scientific terms brought mental processes within the scope of objective analysis. In this context, the notion of resemblance —as a causal notion referring to the process of construction— may play a key role in an objective account of mental processes and in revealing the reason why different representations in different systems have the same reference.

The developmental conception of inheritance defended by Agential Teleosemantics puts the emphasis on a particular aspect of the study of inheritance: parent-offspring interactions during cognitive development. If we want "to take development seriously" (Robert, 2004) and avoid any sort of preformationsm or vacuous explanations, we should look into the dynamics of development coupled with the environmental context of development. Indeed, within DST, eco-devo, and niche construction, *parental effects* are a central research area that ranges

from animal behavior to the study of plants (Badyaev & Uller, 2009; Klopfer, 2001; Uller, 2008). In this context, understanding cognitive development requires analyzing "the interactive unit, parent/child, which is examined within the context of both biological and social/environmental constraints" (Klopfer, 2001, 168).

In general, looking into trans-generational dynamics during cognitive development may help to unravel different pieces of biological information that are central to the acquisition of mental representations. This should be relevant for different cognitive capacities, from primitive perceptual representations to complex ones, such as meta-representational capacities or social skills. In this sense, it is also central to analyze the different channels through which biological information is transmitted during cognitive development: family, society, institutions, and so on.

Developmentally homologous representations

The notion of developmental homology may be relevant in cognition for different reasons. First, in opposition to the historical conception of homology, the developmental conception of homology is committed to the search for developmental causes. The comparative analysis of different species and of their developmental trajectories may nurture the understanding of the developmental trajectory of other species.

Moreover, homology —as it was presented in Section 5.1.2— is connected to the issue of trait characterization. Cognitive science adopts principally a functionalist approach: sameness is based on function. Two cognitive capacities are the same if they *realize* the same function (Balari & Lorenzo, 2014a). However, developmental homology suggests an alternative view. Sameness is established on the basis of developmental history; shared developmental pathways give rise to homologous representations. This is a totally unexplored area of research and, certainly, one facing many difficulties given the complexity of developmental processes. Homologous representations fit perfectly into the framework of Agential Teleosemantics. As content is determined by developmental processes, and a representation is individuated by its content, then similar developmental processes shall lead to representations with the same content.

Cognitive scaffolding

Finally, another conceptual and methodological tool from which cognitive science may benefited from is the notion of scaffolding. Scaffolding processes involve distinct traits that play the role of facilitating further development, as development moves on, these traits may disappear or become hybridized into the system. This phenomenon may be relevant for cognitive development too. Indeed, the acquisition of knowledge and the development of further representational capacities

usually depends on the previous repertoire of representations. The representations and the knowledge acquired at one developmental stage may be essential for the acquisition of more complex representational capacities. This is, indeed, the idea defended in section 9.1.2, where primitive representations may be seen as scaffolds of complex representational capacities.

While scaffolding has proven to be ubiquitous in nature, as argued by Griesemer (2014), research in the dynamics of scaffolding in cognitive development, initiated by Vygotsky (1978), may help uncover the intrinsic dynamics of development through time, as well as the role of external scaffolds, such as parental relationships, society or institutions —cf. Balari and Lorenzo (2018); Bickhard (2005); Hoffmeyer (2008b); Korbak (2015); Sterelny (2010) for some examples of the use of the concept of scaffolding in cognitive development.

9.4 Summary

This chapter is just an attempt to picture how the naturalization of intentionality would be achieved if we adopt Agential Teleosemantics. Further analysis, case studies, and discussion are needed in order to get a full-fledged proposal and to connect the proposed naturalization of intentionality with other problems traditionally associated with it. In Section 9.1 I focused principally on two central aspects for any teleosemantic project: how normative aboutness is defined and how content is determined.

Concerning the first issue, the characterization of teleological development presented in Chapter 8 comes to the fore to understand the normativity of mental representations. In this path, the first step was to distinguish between the two ways of being in cognitive systems; as with any other trait, cognitive systems are involved in processes of change and processes of maintenance. The different ways of being entail, therefore, distinct kinds of goals and norms: norms and goals about the adaptivity of developmental processes, and norms and goals about the adaptivity of physiological processes. Once this distinction was traced, I argued that cognitive development is the process that defines the goals and norms of mental representations insofar as developmental processes are responsible for constructing the constitutive dimension of cognitive systems. At any ontogenetic stage in the lifespan of an organism, the norms and goals of a system were constructed by a teleological ontogenetic process. In conclusion, in opposition to etiological teleosemantics, Agential Teleosemantics states that *a representation must represent whatever it was developed to represent.*

Concerning the issues of content determination, I started by noting that the notion of determination should be taken from Crick's view of information: the developmental resources that provide causal specificity in development. I distinguished

between three kinds of biological information that participate in representational development and show invariance relations with representational content. First, there are different sources of biological information involved in the development of the anatomy and physiology of representational systems. Second, in the case of primitive representations, I introduced the notion of environmental-perceptual biological information (EPBI). EPBI concerns those features of the environment that interact during development with the perceptual system in a way that the representation obtained would refer to such features. Finally, I argued that primitive representations themselves operate as biological information involved in the development of complex representational capacities. This classification of biological information contributes to arguing that no semantic information would be needed in the construction of primitive representations. I supported the idea that once a pinch of underived intentionality is naturalized, such a pinch may work as the building blocks in the construction of complex representational capacities in further ontogenetic stages. In this vein, my proposal is aligned with the referentialist underpinnings of teleosemantics insofar as primitive representations concern those representations whose content is not derived from other representations.

I opened Section 9.2 by stating similitude between developmental explanations and cognitive explanations, namely that both kinds of explanations adopt some sort of informational teleosemantics, where normative aboutness is attributed to a system to understand how the system achieves a goal. I continued by introducing different open questions concerning the co-extensionality of mind, cognition, and agency. My aim was not to defend any clear position; rather, I limited myself to argue that the Life-Mind Continuity Thesis may be approached by investigating the different properties of representational systems found in nature. While I argued that representational systems are present across the living realm, I did not argue for the idea that mental or psychological properties should be extended beyond cognitive systems. In this sense, a classical (and hard) question is left open for further work concerning the difference between representational systems in development and cognitive science.

I ended this chapter with some programmatic comments concerning the possibility of extending the DT to the cognitive domain. The motivation for this project comes from the fact that most theories in the cognitive sciences that are motivated by or grounded in biology adopt a neo-Darwinian framework. I began by noting the old history of dichotomic thinking that hosts contemporary dichotomies in cognitive science. Against the traditional debate between empiricists and nativists, the alternative is a constructivist view of representations, which, rather than asking where representations come from, attempts to search for the different causes and processes that interact during development to construct mental representations. Similarly, the idea of birth as the 'birth of experience' should be abandoned once

we look into pre-natal development and appreciate the importance of intrauterine (or 'intraoval') experiences in cognitive development. These remarks toward a non-dichotomous view of development are motivated and justified by different areas in the DT, such as eco-devo, niche construction, embryology, and developmental psychobiology. I closed Section 9.3 by pointing out that other conceptual approaches and methodologies coming from the DT may be incorporated into cognitive science, such as cross-generational interactions during development, the search for homologous (developmental) representations, and the role of scaffolding and hybridization processes in development.

Conclusions

We have reached the end of this thesis. The principal aim was to analyze different proposals around two riddles in the Life Sciences: Kant's Puzzle and Brentano's Problem. This aim was fulfilled by providing a three-step analysis: the presentation of mainstream answers to the riddles, the critical evaluation of these mainstream answers, and finally proposing alternative answers. As is usually the case in philosophy, each of these steps allows for the opportunity for further analysis of important issues related to the problems treated here. As is usually the case in the philosophy of science, theoretical approaches, such as the one provided here, may be connected with different empirical evidence and suggest new research areas. While my discussion is primarily theoretical, the analysis of specific cases could have been of some help to elucidate the ideas exposed here, but the aim of offering an interdisciplinary analysis of the theoretical foundations of biology and cognitive science left no space for an accurate treatment of empirical cases. Connecting my theoretical reflections with the empirical domain would clearly enrich the ideas exposed in these pages and also may contribute to developing further ideas around Agential Teleosemantics.

Causes and abstractions

Engels' words in the epigraph at the opening of this thesis single out the central reason why mainstream answers deserve to be revised: the theory of natural selection built around the MS is not a theory about the causes of evolution. It is a theory about the effects of evolutionary causes in populations, about how "the effects become fixed" (Engels, 1878, 82-83). The DT has taken over in the quest for the causes of aptness. While I stressed that the DT is not a solid and integrated biological theory yet, the common idea within different areas and theories within the DT is that individual-level processes are central in the explanation of evolutionary causes. The most drastic version is supported by the Statisticalist School: all causes of adaptive evolution are ontogenetic.

In the introduction of this thesis, I stated that the argumental logic to be adopted is that of a conditional: what theoretical consequences for the Life Sciences can we recognize *if* we adopt a DT framewrok? Particularly, *if* we adopt

a DT framework, how Kant's Puzzle and Brentano's Problem should be solved? If Engels' quote illustrates the core of the DT, the first thing that we say about teleology and intentionality is that mainstream answers deserve to be challenged. The issue turns around causality. Kant's Puzzle and Brentano's Problem arise respectively when we search for the causes of teleological and intentional explanations. A first conclusion, argued throughout the second part of the thesis, is that *the MS framework is not a theory about the causes of aptness, and consequently, it is not appropriate for naturalizing teleology, nor for anchoring a teleosemantic project.*

The developmental scale

Beyond criticism, this thesis aims to state some programmatic ideas motivated by the DT. If causes are not be found at the populational level, we should look at the individual one. Part III analyzes the physiological ad developmental scales in the search for the causes of aptness and, consequently, the possibility for naturalizing teleology and intentionality. However, when confronting these issues, I highlighted a difference between the physiological and developmental scales: non-orthodox contemporary views on natural teleology are principally concerned with physiological processes. Teleological development stills requires a proper treatment. So, while I presented autonomous systems theory —and its many forerunners in cybernetics and systems biology— as a suitable account for teleological physiology and organismal agency, it seems that the issue of teleological development is not addressed properly in the contemporary philosophy of biology.

Moreover, there is a fundamental reason why a developmental approach to teleology is also necessary: not all is about self-maintenance. Self-maintenance is a necessary but not sufficient phenomenon to understand the goal-directedness of an organism's activity. This was introduced by stressing the double way of being of organisms: beyond maintenance, organisms also change. Organismal agency, an individual-level phenomenon, is devoted to fulfilling both goals throughout the lifespan of an organism: maintaining the system and changing it. While teleological physiology is about self-maintenance, a developmental perspective incorporates the chaining nature of living systems and how this goal regulates the construction of complex and apt living systems.

Agential Teleosemantics

The view on teleological development presented here reinvigorates a number of ideas introduced throughout the history of biology. In the path of granting organisms the responsibility for creating order and adaptive complexity, the first step is to acknowledge the importance of avoiding any sort of preformationism. Here

information starts to be the main protagonist. Against preformationist views of information, an order-from-order perspective must defend where information is created in ontogeny, and where order does not preexist developmental processes. Other specific forerunners also contribute to forging the view presented here. Von Baer's terminological insights and his emphasis on harmonic regulations during ontogeny become central to understanding the goal-directedness of organismal activity as the *proper* force that produces aptness. Also, von Uexküll's contribution is integrated into Agential Teleosemantics: first, the idea that an organism's activity is relative to the organism's experiences —an idea that echoes Lewontin's constructivist view; and second, the importance of understanding goal-directedness not as an adult-oriented process but as a process of construction relative to the specific ontogenetic stage where the developmental system is located. Developmental organicists also contributed to the path toward teleological development. The main lesson that we can reach from them is that developmental processes have a degree of freedom, and freedom enables creativity: the capacity of regulating developmental processes for adaptive reasons. All the elements introduced by different forerunners are incorporated in Walsh's Agential Perspective.

However, Walsh's Agential Perspective is not explicitly committed to a non-adult-oriented view of goals, nor is it explicit enough about the double way of being of organisms and their respective goals. These, surely, are important ingredients to add to Walsh's view on teleological development. However, the main shortcoming in Walsh's proposal is that he provides a misguided analysis of invariance relations in teleological explanations. The analysis of invariance relations in teleological explanations, rather than leading to the conclusion that teleological explanations are not casual, as Walsh argues, reveals the proper causal structure of teleological explanations: teleological explanations are causal explanations insofar as a system *having a goal* causes certain *means* that are directed to the fulfillment of the goal. Invariance relations arise between a system having a goal and the means toward the goal, where the former causes the latter. The recognition of the causal structure of teleological explanations not only demands some rearrangements of Walsh's proposal but also enables me to defend a cardinal idea in this thesis: that teleological explanations in development have the same explanatory structure as intentional explanations. Agential Teleosemantics attempts to provide a non-preformationist view of goals and to explain the changing nature of developmental processes, but it is primarily a theory that explains how teleological explanations can be understood in intentional terms.

The key terms to understand teleological development are two: information and agency. Agential Teleosemantics integrates different approaches to development on the basis of informational processes and distributed specificity with the view of organismal agency defended by the DT. Intentionality arises in develop-

mental processes once we recognize its two central ingredients: *aboutness*, arising by informational processes informing about the conditions of development, and *normativity*, as the result of agents regulating such conditions in an adaptively directed way. Having a goal is about a system interpreting and using biological information to produce a certain means that conduces the system towards the goal.

What is the goal of a developmental system? I defined developmental goals as the aim of producing adaptive changes at each ontogenetic stage in the constitutive dimension. The emphasis on *change* captures the changing way of being of living organisms beyond maintenance. The emphasis on changing the *constitutive* dimension accounts for the fact that developmental processes generate and create adaptive complex systems: the parts of a system are the result of developmental processes. Finally, stressing that goals are always relative to *ontogenetic scales* enables me to avoid any sort of preformationism: goals are not adult-oriented, they do not pre-exist developmental processes. Rather, developmental goals are always relative to the current conditions of a developmental system.

Agential Teleosemantics in Evolutionary Theory

Agential Teleosemantics comes to occupy a place within the DT that is mandatory to fill once populational forces are evaporated; i.e. once the MS is not suitable for dealing with the causes of aptness and, consequently, neither for naturalizing teleology. If the cross-generational resemblance needed in natural selection processes is not exclusively the product of genetic inheritance, then something else is needed. Enter developmental conceptions of inheritance and extended systems of inheritance. If (heritable) phenotypic variations cannot be fully understood just by looking at genetic drift and genetic mutation, then we need an alternative view of variation. Enter the plasticity-first view on phenotypic variation and self-organization processes. Finally, if individual fitness cannot be traced back to the genetic underpinning of an individual, then something else must explain it. Enter agency, and the organismal capacity of producing and maintaining apt systems. In sum, if genes do not produce order, something else must do it. Enter organisms. Agential Teleosemantics integrates all these insights from the DT into a naturalist view of teleological development.

Another way of setting the scene is by noting that the abandonment of populational causes requires an explanation of aptness based on individual causation. Individual processes cause aptness, the Statisticalist School supports, and evolution is about the populational and historical consequences of what organisms do. Agential Teleosemantics attempts to account for such requirement and provides a view of individual-level processes that produce apt traits by pointing out how inner processes are responsible for generating complex and adaptive organisms. Rather

than the externalist view where natural selection is seen as the directive force in evolution, organisms themselves and their inner distinctive qualities become the source of evolutionary causation.

As stressed in the introduction of this thesis, my principal aim is to address the problems of intentionality and teleology assuming the DT framework. One of the main conclusions is that central explanatory roles in theoretical biology should be transferred from the populational to the individual level. This, nonetheless, does not mean that the populational perspective adopted by the MS should be abandoned at all. Certainly, I did not get into a detailed scrutiny of whether the MS deserves modification, extension, or revolution. But I did support a pluralist view of biological explanations motivated by the Statisticalist School. Accordingly, different levels of analysis (the individual and the populational ones) account for different phenomena using different explanatory strategies. It is therefore a central task in the contemporary philosophy of biology to assess such epistemic pluralism; or, in other words, to evaluate how the DT and the MS can coherently be brought together in the quest for a better and streamlined view of biology.

Bibliography

Aaby, B. H., & Desmond, H. (2021). Niche construction and teleology: Organisms as agents and contributors in ecology, development, and evolution. *Biology & Philosophy*, *36*(5), 1–20. doi: 10.1007/s10539-021-09821-2

Abrams, M. (2012). Measured, modeled, and causal conceptions of fitness. *Frontiers in Genetics (Evolutionary and Population Genetics)*, *3*(196), 1–12. doi: 10.3389/fgene.2012.00196

Ågren, J. A. (2021). *The Gene's-Eye View of Evolution*. Oxford: Oxford University Press.

Alberch, P. (1980). Ontogenesis and morphological diversification. *American Zoologist*, *20*(4), 653–667. doi: 10.1093/icb/20.4.653

Alberch, P. (1982). Developmental constraints in evolutionary processes. In J. T. Bonner (Ed.), *Evolution and Development. Dahlem Workshop Reports* (Vol. 22, pp. 313–332). Berlin: Springer. doi: 10.1007/978-3-642-45532-2_15

Alberch, P. (1989). The logic of monsters: Evidence for internal constraint in development and evolution. *Geobios*, *22*(Supplement 2), 21–57. doi: 10.1016/S0016-6995(89)80006-3

Alberch, P. (1991). From genes to phenotype: Dynamical systems and evolvability. *Genetica*, *84*(1), 5–11. doi: 10.1007/bf00123979

Amundson, R. (1988). Logical adaptationism. *Behavioral and Brain Sciences*, *11*(3), 505–506. doi: 10.1017/S0140525X00058623s

Amundson, R. (1989). The trials and tribulations of selectionist explanations. In K. Hahlweg & C. A. Hooker (Eds.), *Issues in Evolutionary Epistemology: Contemporary Engagements between Analytic and Continental Thought* (pp. 413–432). Albany: SUNY Press.

Amundson, R. (1990). Doctor Dennett and Doctor Pangloss: Perfection and selection in biology and psychology. *Behavioral and Brain Sciences*, *13*(3), 577–581. doi: 10.1017/S0140525X00080237

Amundson, R. (1994). Two concepts of constraint: Adaptationism and the challenge from developmental biology. *Philosophy of Science*, *61*(4), 556–578. doi: 10.1086/289822

Amundson, R. (2000). Against normal function. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *31*(1), 33–53. doi: 10.1016/S1369-8486(99)00033-3

Amundson, R. (2005). *The Changing Role of the Embryo in Evolutionary Thought: Roots of Evo-Devo*. Cambridge: Cambridge University Press.

Amundson, R. (2006). EvoDevo as cognitive psychology. *Biological Theory*, *1*(1), 10–11. doi: 10.1162/biot.2006.1.1.10

Amundson, R., & Lauder, G. V. (1994). Function without purpose: The uses of causal role function in evolutionary biology. *Biology & Philosophy*, *9*(4), 443–469. doi: 10.1007/BF00850375

Anderson, P. W. (1972). More is different: Broken symmetry and the nature of the hierarchical structure of science. *Science*, *177*(4047), 393–396. doi: 10.1126/science.177.4047.393

Appel, T. (1987). *The Cuvier-Geoffroy Debate: French Biology in the Decades before Darwin*. New York: Oxford University Press.

Ariew, A. (2003). Ernst Mayr's 'ultimate/proximate' distinction reconsidered and reconstructed. *Biology & Philosophy*, *18*(4), 553–565. doi: 10.1023/A:1025565119032

Ariew, A., Cummins, R., & Perlman, M. (Eds.). (2002). *Functions: New Essays in the Philosophy of Psychology and Biology*. Oxford: Oxford University Press.

Ariew, A., & Lewontin, R. C. (2004). The confusions of fitness. *The British Journal for the Philosophy of Science*, *55*(2), 347–363. doi: 10.1093/bjps/55.2.347

Ariew, A., Rice, C., & Rohwer, Y. (2015). Autonomous-statistical explanations and natural selection. *The British Journal for the Philosophy of Science*, *66*(3), 635–658. doi: 10.1093/bjps/axt054

Arnellos, A., Spyrou, T., & Darzentas, J. (2010). Towards the naturalization of agency based on an interactivist account of autonomy. *New Ideas in Psychology*, *28*(3), 296–311. doi: 10.1016/j.newideapsych.2009.09.005

Artiga, M. (2013). Reliable misrepresentation and teleosemantics. *Disputatio*, *V*(37). doi: 10.2478/disp-2013-0020

Artiga, M. (2014a). Signaling without cooperation. *Biology & Philosophy*, *29*(3), 357–378. doi: 10.1007/s10539-014-9436-0

Artiga, M. (2014b). Teleosemantics and pushmi-pullyu representations. *Erkenntnis*, *79*(3), 545–566. doi: 10.1007/s10670-013-9517-5

Artiga, M. (2020). Models, information and meaning. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *82*(5), 101284. doi: https://doi.org/10.1016/j.shpsc.2020.101284

Artiga, M. (2021). Biological functions and natural selection: A reappraisal.

European Journal for Philosophy of Science, *11*(2), 54. doi: 10.1007/s13194-021-00357-6

Artiga, M., & Martínez, M. (2016). The organizational account of function is an etiological account of function. *Acta Biotheoretica*, *64*(2), 105–117. doi: 10.1007/s10441-015-9256-x

Ashby, W. R. (1956). *An Introduction to Cybernetics*. London: Chapman & Hall Ltd.

Ashby, W. R. (1991). Principles of the self-organizing system. In G. J. Klir (Ed.), *Facets of Systems Science* (pp. 521–536). Boston: Springer.

Avital, E., & Jablonka, E. (2000). *Animal Traditions: Behavioural Inheritance in Evolution*. Cambridge: Cambridge University Press.

Ayala, F. J. (1970). Teleological explanations in evolutionary biology. *Philosophy of Science*, *37*(1), 1–15. doi: 10.1086/288276

Ayala, F. J. (2007). Darwin's greatest discovery: Design without designer. *Proceedings of the National Academy of Sciences*, *104*(Supplement 1), 8567–8573. doi: 10.1073/pnas.0701072104

Babcock, G., & McShea, D. W. (2021). An externalist teleology. *Synthese*, *199*(3), 8755–8780. doi: 10.1007/s11229-021-03181-w

Badyaev, A. V., & Uller, T. (2009). Parental effects in ecology and evolution: mechanisms, processes and implications. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1520), 1169–1177. doi: 10.1098/rstb.2008.0302

Baedke, J. (2013). The epigenetic landscape in the course of time: Conrad Hal Waddington's methodological impact on the life sciences. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *44*(4), 756–773. doi: 10.1016/j.shpsc.2013.06.001

Baedke, J. (2018a). *Above the Gene, Beyond Biology: Toward a Philosophy of Epigenetics*. Pittsburgh, PA: University of Pittsburgh Press.

Baedke, J. (2018b). O organism, where art thou? Old and new challenges for organism-centered biology. *Journal of the History of Biology*, *52*(2), 293–324. doi: 10.1007/s10739-018-9549-4

Baedke, J., & Gilbert, S. F. (2021). Evolution and development. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2021 ed.). Metaphysics Research Lab, Stanford University. `https://plato.stanford.edu/archives/fall2021/entries/evolution-development/`.

von Baer, K. E. (1886a). Über den Zweck in den Vorgängen der Natur — Erste Hälfte. Über Zweckmäßigkeit oder Zielstrebigkeit überhaupt. In *Reden gehalten in wissenschaftlichen Versammlungen und kleinere Aufsätze vermischten Inhalts. Zweiter Theil. Studien aus dem Gebiete der Naturwissenschaften*

(pp. 49–106). Braunschweig: Friedrich Vieweg und Sohn.

von Baer, K. E. (1886b). Über Zielstrebigkeit in den organischen Körpern insbesondere. In *Reden gehalten in wissenschaftlichen Versammlungen und kleinere Aufsätze vermischten Inhalts. Zweiter Theil. Studien aus dem Gebiete der Naturwissenschaften* (pp. 170–234). Braunschweig: Friedrich Vieweg und Sohn.

Balari, S., & Lorenzo, G. (2010). ¿Para qué sirve un ballestrinque? Reflexiones sobre el funcionamiento de artefactos y organismos en un mundo sin funciones. *Teorema: Revista internacional de filosofía*, *29*(3), 57–76.

Balari, S., & Lorenzo, G. (2013). *Computational Phenotypes: Towards an Evolutionary Developmental Biolinguistics*. Oxford: Oxford University Press.

Balari, S., & Lorenzo, G. (2014a). Ahistorical homology and multiple realizability. *Philosophical Psychology*, *28*(6), 881–902. doi: 10.1080/09515089.2014.949004

Balari, S., & Lorenzo, G. (2014b). The end of development. *Biological Theory*, *10*(1), 60–72. doi: 10.1007/s13752-014-0180-0

Balari, S., & Lorenzo, G. (2015). It is an organ, it is new, but it is not a new organ. Conceptualizing language from a homological perspective. *Frontiers in Ecology and Evolution*, *3*, 58. doi: 10.3389/fevo.2015.00058

Balari, S., & Lorenzo, G. (2018). The internal, the external and the hybrid: The state of the art and a new characterization of language as a natural object. *Glossa. A Journal of General Linguistics*, *3*(1), 1–3. doi: 10.5334/gjgl.330

Balari, S., & Lorenzo, G. (2021). Evo-devo of language and cognition. In L. Nuño de la Rosa & G. B. Müller (Eds.), *Evolutionary Developmental Biology: A Reference Guide* (pp. 1221–1233). Cham: Springer Nature.

Balari, S., Lorenzo, G., & Sultan, S. E. (2020). Language acquisition and ecodevo processes: The case of the lexicon-syntax interface. *Biological Theory*, *15*(3), 148–160. doi: 10.1007/s13752-020-00352-9

Baldwin, J. M. (1896). A new factor in evolution. *The American Naturalist*, *30*(354), 441–451. doi: 10.1086/276408

Barandiaran, X. (2008). *Mental life. A Naturalized Approach to the Autonomy of Cognitive Agents* (Unpublished doctoral dissertation). Universidad del País Vasco — Euskal Herriko Unibertsitatea.

Barandiaran, X., Di Paolo, E., & Rohde, M. (2009). Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action. *Adaptive Behavior*, *17*(5), 367–386. doi: 10.1177/1059712309343819

Barandiaran, X., & Egbert, M. (2014). Norm-establishing and norm-following in autonomous agency. *Artificial Life*, *20*(1), 5–28. doi: 10.1162/artl_a_00094

Barandiaran, X., & Moreno, A. (2008). Adaptivity: From metabolism to behavior. *Adaptive Behavior*, *16*(5), 325–344. doi: 10.1177/1059712308093868

Barbieri, M. (2014). From biosemiotics to code biology. *Biological Theory*, *9*(2), 239–249. doi: 10.1007/s13752-013-0155-6

Barbieri, M. (2015). *Code Biology. A New Science of Life*. Cham: Springer.

Bard, J. B. L. (2011). A systems biology representation of developmental anatomy. *Journal of Anatomy*, *218*(6), 591–599. doi: 10.1111/j.1469-7580.2011.01371.x

Bard, J. B. L. (2013). Driving developmental and evolutionary change: A systems biology view. *Progress in Biophysics and Molecular Biology*, *111*(2-3), 83–91. doi: 10.1016/j.pbiomolbio.2012.09.006

Barkow, J., Cosmides, L., & Tooby, J. (1992). *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. New York: Oxford University Press.

Bateson, P. (2005). The return of the whole organism. *Journal of Biosciences*, *30*(1), 31–39. doi: 10.1007/bf02705148

Bateson, P., & Gluckman, P. (2011). *Plasticity, Robustness, Development and Evolution*. Cambridge: Cambridge University Press.

Bateson, P., & Mameli, M. (2007). The innate and the acquired: Useful clusters or a residual distinction from folk biology? *Developmental Psychobiology*, *49*(8), 818–831. doi: 10.1002/dev.20277

Beatty, J. (2016). The creativity of natural selection? Part I: Darwin, Darwinism, and the Mutationists. *Journal of the History of Biology*, *49*(4), 659–684. doi: 10.1007/s10739-016-9456-5

Beatty, J. (2019). The creativity of natural selection? Part II: The synthesis and since. *Journal of the History of Biology*, *52*(4), 705–731. doi: 10.1007/s10739-019-09583-4

Bechtel, W., & Bich, L. (2021). Grounding cognition: Heterarchical control mechanisms in biology. *Philosophical Transactions of the Royal Society B*, *376*(1820), 20190751. doi: 10.1098/rstb.2019.0751

Bergson, H. (1907). *L'Évolution créatrice*. Paris: Les Presses Universitaires de France.

Bermúdez, J. L. (2014). *Cognitive Science: An Introduction to the Science of the Mind*. Cambridge: Cambridge University Press.

Bernard, C. (1879). *Leçons sur les phénomènes de la vie commune aux animaux et aux végétaux* (Vol. 2). Paris: Baillière.

von Bertalanffy, L. (1933). *Modern Theories of Development: An Introduction to Theoretical Biology*. Oxford: Oxford University Press.

von Bertalanffy, L. (1950). The theory of open systems in physics and biology. *Science*, *111*(2872), 23–29. doi: 10.1126/science.111.2872.23

von Bertalanffy, L. (1952). *Problems of Life: An Evaluation of Modern Biological Thought*. New York: Wiley.

von Bertalanffy, L. (1969). *General System Theory: Foundations, Development, Applications*. New York: George Braziller.

von Bertalanffy, L. (1972). The model of open systems: Beyond molecular biology. In A. D. Breck & W. Yourgrau (Eds.), *Biology, History, and Natural Philosophy* (pp. 17–30). New York: Springer. doi: 10.1007/978-1-4684-1965-8_2

Bertolotti, T., & Magnani, L. (2017). Theoretical considerations on cognitive niche construction. *Synthese*, *194*(12), 4757–4779. doi: 10.1007/s11229-016-1165-2

Berwick, R. C., & Chomsky, N. (2016). *Why Only Us: Language and Evolution*. Cambridge, MA: The MIT Press.

Bich, L., Mossio, M., Ruiz-Mirazo, K., & Moreno, . (2016). Biological regulation: Controlling the system from within. *Biology & Philosophy*, *31*(2), 237–265. doi: 10.1007/s10539-015-9497-8

Bichat, X. (1805). *Recherches physiologiques sur la vie et la mort* (3rd ed.). Paris: Brosson.

Bickhard, M. H. (2000a). Autonomy, function, and representation. *Communication and Cognition—Artificial Intelligence*, *17*(3-4), 111–131.

Bickhard, M. H. (2000b). Information and representation in autonomous agents. *Cognitive Systems Research*, *1*(2), 65–75. doi: 10.1016/s1389-0417(99)00007-8

Bickhard, M. H. (2003). The biological emergence of representation. In T. Brown & L. Smith (Eds.), *Reductionism and the Development of Knowledge* (pp. 115–142). New York: Lawrence Erlbaum.

Bickhard, M. H. (2004). The dynamic emergence of representation. In H. Clapin, P. Staines, & P. Slezak (Eds.), *Representation in Mind: New Approaches to Mental Representation* (pp. 71–90). Amsterdam: Elsevier. doi: 10.1016/B978-008044394-2/50007-5

Bickhard, M. H. (2005). Functional scaffolding and self-scaffolding. *New Ideas in Psychology*, *23*(3), 166–173. doi: 10.1016/j.newideapsych.2006.04.001

Bickhard, M. H. (2009a). The biological foundations of cognitive science. *New Ideas in Psychology*, *27*(1), 75–84. doi: 10.1016/j.newideapsych.2008.04.001

Bickhard, M. H. (2009b). Interactivism: A manifesto. *New Ideas in Psychology*, *27*(1), 85–95. doi: 10.1016/j.newideapsych.2008.05.001

Birch, J. (2014). Propositional content in signalling systems. *Philosophical Studies*, *171*(3), 493–512. doi: 10.1007/s11098-014-0280-5

Birch, J., Martínez, M., & Artiga, M. (Eds.). (2020). *Signaling Systems. Special Issue of Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* (Vol. 84 December).

Bock, W. J., & von Wahlert, G. (1965). Adaptation and the Form-Function

Complex. *Evolution*, *19*(3), 269–299. doi: 10.2307/2406439

Boogerd, F. (2007). *Systems Biology: Philosophical Foundations*. Amsterdam: Elsevier.

Boyd, R., & Richerson, P. J. (1985). *Culture and the Evolutionary Process*. Chicago, IL: The University of Chicago Press.

Brentano, F. (1995). *Psychology from an Empirical Standpoint*. London New York: Routledge.

Brigandt, I. (2002). Homology and the origin of correspondence. *Biology & Philosophy*, *17*(3), 389–407. doi: 10.1023/a:1020196124917

Brigandt, I. (2013). Explanation in biology: Reduction, pluralism, and explanatory aims. *Science & Education*, *22*(1), 69–91. doi: 10.1007/s11191-011-9350-7

Brigandt, I., & Griffiths, P. E. (2007). The importance of homology for biology and philosophy. *Biology & Philosophy*, *22*(5), 633–641. doi: 10.1007/s10539-007-9094-6

Bromberger, S. (1966). Why-questions. In R. G. Colodny (Ed.), *Mind and Cosmos* (p. 86-111). Pittsburgh: University of Pittsburgh Press.

Buss, D. (2008). *Evolutionary Psychology: the New Science of the Mind* (3rd ed.). Boston: Pearson.

Calcott, B., Pocheville, A., & Griffiths, P. E. (2020). Signals that make a difference. *The British Journal for the Philosophy of Science*, *71*(1), 233–258. doi: 10.1093/bjps/axx022

Camazine, S., Deneubourg, J.-L., Franks, N. R., Sneyd, J., Bonabeau, E., & Theraula, G. (2003). *Self-Organization in Biological Systems*. Princeton, NJ: Princeton University Press.

Canguilhem, G. (2012). *On the Normal and the Pathological*. Dordrecht: Springer.

Cannon, W. B. (1929). Organization for physiological homeostasis. *Physiological Reviews*, *9*(3), 399–431. doi: 10.1152/physrev.1929.9.3.399

Caporael, L. R. (2001). Evolutionary psychology: Toward a unifying theory and a hybrid science. *Annual Review of Psychology*, *52*, 607–628. doi: 10.1146/annurev.psych.52.1.607

Carey, S. (2009). *The Origin of Concepts*. Oxford: Oxford University Press.

Carroll, S. B. (2005). *Endless Forms Most Beautiful: The New Science of Evo Devo and the Making of the Animal Kingdom*. New York: W. W. Norton & Co.

Carruthers, P., & Laurence, S. (Eds.). (2005). *The Innate Mind: Structure and Contents* (Vol. 1). Oxford: Oxford University Press.

Carruthers, P., & Laurence, S. (Eds.). (2007a). *The Innate Mind: Culture and Cognition* (Vol. 2). Oxford: Oxford University Press Inc.

Carruthers, P., & Laurence, S. (Eds.). (2007b). *The Innate Mind: Foundations and the Future* (Vol. 3). Oxford: Oxford University Press.

Casal, J. (2004). Signalling for developmental plasticity. *Trends in Plant Science*, *9*(6), 309–314. doi: 10.1016/j.tplants.2004.04.007

Chemero, A. (1998). Teleosemantics and the critique of adaptationism. *Evolution and Cognition*, *4*(2), 136–144.

Chiu, L. (2019). Decoupling, commingling, and the evolutionary significance of experiential niche construction. In T. Uller & K. N. Laland (Eds.), *Evolutionary Causation: Biological and Philosophical Reflections* (pp. 299–321). Cambridge, MA: The MIT Press.

Chiu, L., & Gilbert, S. F. (2015). The birth of the holobiont: Multi-species birthing through mutual scaffolding and niche construction. *Biosemiotics*, *8*(2), 191–210. doi: 10.1007/s12304-015-9232-5

Chomsky, N. (2000). *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.

Christensen, W., & Bickhard, M. H. (2002). The process dynamics of normative function. *The Monist*, *85*(1), 3–28. doi: 10.2307/27903755

Churchland, P. M. (1988). *Matter and Consciousness: A Contemporary Introduction to the Philosophy of Mind.* Cambridge, MA: The MIT Press.

Cowie, F. (1999). *What's Within? Nativism Reconsidered*. Oxford: Oxford University Press.

Craver, C. F. (2007). *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford: Oxford University Press.

Craver, C. F. (2013). Functions and mechanisms: A perspectivalist view. In P. Huneman (Ed.), *Functions: Selection and Mechanisms* (pp. 133–158). Dordrecht: Springer.

Crick, F. H. C. (1958). On protein synthesis. In F. K. Sanders (Ed.), *The Biological Replication of Macromolecules. Symposia of the Society for Experimental Biology* (Vol. 12, pp. 138–163). Cambridge: Cambridge University Press.

Crick, F. H. C. (1970). Central dogma of molecular biology. *Nature*, *227*(5258), 561–563. doi: 10.1038/227561a0

Cummins, R. (1975). Functional analysis. *The Journal of Philosophy*, *72*(20), 741–765. doi: 10.2307/2024640

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray.

Davidson, D. (1987). Knowing one's own mind. *Proceedings and Addresses of the American Philosophical Association*, *60*(3), 441–458. doi: 10.2307/3131782

Davies, P. S. (2000). Malfunctions. *Biology & Philosophy*, *15*(1), 19–38. doi: 10.1023/a:1006525318699

Dawkins, R. (1976). *The Selfish Gene*. Oxford: Oxford University Press.

Dawkins, R. (1978). Replicator selection and the extended phenotype. *Zeitschrift*

für Tierpsychologie, *47*(1), 61–76. doi: 10.1111/j.1439-0310.1978.tb01823.x

Dawkins, R. (1982). *The Extended Phenotype*. Oxford: Oxford University Press.

Dawkins, R. (1994). Burying the vehicle. *Behavioral and Brain Sciences*, *17*(4), 616–617. doi: 10.1017/S0140525X00036207

Dehaene, S. (1997). *The Number Sense: How the Mind Creates Mathematics*. New York: Oxford University Press.

Delage, Y. (1903). *L'Hérédité et les grands problèmes de la biologie générale*. Paris: C. Reinwald.

de Llanza Varona, M. (2022). *Synergistic Information Requires Computation: Against Anti-representationalism* (Unpublished master's thesis). Universitat de Barcelona, Barcelona.

Dennett, D. C. (1987). *The Intentional Stance*. Cambridge, MA: The MIT Press.

Dennett, D. C. (1990). The interpretation of texts, people and other artifacts. *Philosophy and Phenomenological Research*, *50*(Supplement), 177–194. doi: 10.2307/2108038

Dennett, D. C. (1994). The myth of original intentionality. In E. Dietrich (Ed.), *Thinking Computers and Virtual Persons: Essays on the Intentionality of Machines* (pp. 91–107). San Diego: Academic Press. doi: 10.1016/B978-0-12-215495-9.50009-0

Dennett, D. C. (1995). *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon & Schuster.

Dennett, D. C. (2017). *From Bacteria to Bach and Back: The Evolution of Minds*. New York: W. W. Norton & Company.

Desmond, H., & Huneman, P. (2020). The ontology of organismic agency: A Kantian approach. In A. Altobrando & P. Biasetti (Eds.), *Natural Born Monads: On the Metaphysics of Organisms and Human Individuals* (pp. 33–64). Berlin: De Gruyter. doi: TheOntologyofOrganismicAgency:A{Kantian}Approach

DeWitt, T. J., & Scheiner, S. M. (Eds.). (2004). *Phenotypic Plasticity: Functional and Conceptual Approaches*. Oxford: Oxford University Press.

DiFrisco, J. (2021). Developmental homology. In L. Nuño de la Rosa & G. B. Müller (Eds.), *Evolutionary Developmental Biology: A Reference Guide* (pp. 85–97). Cham: Springer.

DiFrisco, J., & Jaeger, J. (2020). Genetic causation in complex regulatory systems: An integrative dynamic perspective. *BioEssays*, *42*(6), 1900226. doi: 10.1002/bies.201900226

Di Paolo, E., Barandiaran, X., Beaton, M., & Buhrmann, T. (2014). Learning to perceive in the sensorimotor approach: Piaget's theory of equilibration interpreted dynamically. *Frontiers in Human Neuroscience*, *8*, 551. doi: 10.3389/fnhum.2014.00551

Dobzhansky, T. (1937). *Genetics and the Origin of Species*. New York: Columbia University Press.

Dobzhansky, T. (1955). A review of some fundamental concepts and problems of population genetics. *Cold Spring Harbor Symposia on Quantitative Biology*, *20*, 1–15. doi: 10.1101/SQB.1955.020.01.003

Dobzhansky, T. (1958). Species after Darwin. In S. A. Barnett (Ed.), *A Century of Darwin* (pp. 19–55). London: Heinemann.

Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge, MA: The MIT Press.

Dretske, F. (1988). *Explaining Behavior: Reasons in a World of Causes*. Cambrdige, MA: The MIT Press.

Dretske, F. (2001). Norms, history and the mental. *Royal Institute of Philosophy Supplement*, *49*, 87–104. doi: 10.1017/s1358246100007128

Dretske, F. (2004). Psychological vs. biological explanations of behavior. *Behavior and Philosophy*, *32*(1), 167–177.

Driesch, H. (1908). *The Science and Philosophy of the Organism*. London: Adam and Charles Black. (2 Vols.)

Driesch, H. (1914). *The History & Theory of Vitalism*. London: Macmillan and Co. (Translated from German by C. K. Ogden.)

Dubinin, N. P., & Vaulina, E. N. (1976). The evolutionary role of gravity. *Life Sciences and Space Research*, *14*, 47–55. doi: 10.1515/9783112516843-005

Durham, W. H. (1991). *Coevolution. Genes, Culture, and Human Diversity*. Stanford, CA: Stanford University Press.

Edelmann, J. B., & Denton, M. J. (2006). The uniqueness of biological self-organization: Challenging the Darwinian paradigm. *Biology & Philosophy*, *22*(4), 579–601. doi: 10.1007/s10539-006-9055-5

El-Hani, C. N. (2007). Between the cross and the sword: The crisis of the gene concept. *Genetics and Molecular Biology*, *30*(2), 297–307. doi: 10.1590/S1415-47572007000300001

Emmeche, C., & Kull, K. (Eds.). (2011). *Towards a Semiotic Biology: Life Is the Action of Signs*. London: Imperial Colleague Press.

Endler, J. A. (2020). *Natural Selection in the Wild*. Princeton, NJ: Princeton University Press.

Engels, F. (1878). *Herrn Eugen Dühring's Umwälzung der Wissenschaft*. Leipzig: Genossenschafts-Buchdruckerei. (English translation in Engels, 1976.)

Engels, F. (1976). *Herr Eugen Düring's Revolution in Science (Anti-Duhring)*. New York International Publisher.

Esposito, M. (2013). Heredity, development and evolution: The unmodern synthesis of E.S. Russell. *Theory in Biosciences*, *132*(3), 165–180. doi: 10.1007/s12064-013-0177-4

Etxeberria, A., & Umerez, J. (2006). Organismo y organización en la biología teórica: ¿vuelta al organicismo? *Ludus Vitalis*, *14*(26), 3–38.

Favareau, D. (Ed.). (2010). *Essential Readings in Biosemiotics: Anthology and Commentary*. Dordrecht: Springer.

Feiten, T. E. (2020). Mind after Uexküll: A foray into the worlds of ecological psychologists and enactivists. *Frontiers in Psychology*, *11*, 480. doi: 10.3389/fpsyg.2020.00480

Finlay, B. L. (2007). Endless minds most beautiful. *Developmental Science*, *10*(1), 30–34. doi: 10.1111/j.1467-7687.2007.00560.x

Fodor, J. A. (1975). *The Language of Thought*. Cambridge, MA: Harvard University Press.

Fodor, J. A. (1981). *RePresentations: Philosophical Essays on the Foundations of Cognitive Science*. Cambridge, MA: MIT Press.

Fodor, J. A. (1990). *A Theory of Content and Other Essays*. Cambridge, MA: MIT Press.

Fodor, J. A. (1998). *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press.

Fodor, J. A. (2008). *LOT 2: the Language of Thought Revisited*. Oxford: OUP Oxford.

Fodor, J. A., & Piattelli-Palmarini, M. (2010). *What Darwin Got Wrong*. New York: Farrar, Straus and Giroux.

Fodor, J. A., & Pylyshyn, Z. W. (2015). *Minds without Meanings. An Essay on the Content of Concepts*. Cambridge, MA: The MIT Press.

von Foerster, H. (2003). *Understanding Understanding: Essays on Cybernetics and Cognition*. New York: Springer.

Forgacs, G., & Newman, S. A. (2005). *Biological Physics of the Developing Embryo*. Cambridge: Cambridge University Press.

Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, *100*, 25–50. (English translation in Frege, 1948.)

Frege, G. (1948). Sense and reference. *The Philosophical Review*, *57*(3), 209–230. doi: 10.2307/2181485

Frick, R., Bich, L., & Moreno, A. (2019). An organisational approach to biological communication. *Acta Biotheoretica*, *67*(2), 103–128. doi: 10.1007/s10441-019-09342-2

Futuyma, D. J. (2017). Evolutionary biology today and the call for an extended synthesis. *Interface Focus*, *7*(5), 20160145. doi: 10.1098/rsfs.2016.0145

Gambarotto, A., & Nahas, A. (2022). Teleology and the organism: Kant's controversial legacy for contemporary biology. *Studies in History and Philosophy of Science*, *93*, 47–56. doi: 10.1016/j.shpsa.2022.02.005

Ganson, T. (2018). The senses as signalling systems. *Australasian Journal of*

Philosophy, *96*(3), 519–531. doi: 10.1080/00048402.2017.1381749

García-Carpintero, M., & Macià, J. (Eds.). (2006). *Two-Dimensional Semantics*. Oxford: Oxford University Press.

Gardner, A. (2009). Adaptation as organism design. *Biology Letters*, *5*(6), 861–864. doi: 10.1098/rsbl.2009.0674

Garson, J. (2016). *A Critical Overview of Biological Functions*. Cham: Springer.

Garson, J., & Papineau, D. (2019). Teleosemantics, selection and novel contents. *Biology & Philosophy*, *34*(3), 36. doi: https://doi.org/10.1007/s10539-019-9689-8

Gayon, J. (1998). The concept of individuality in Canguilhem's philosophy of biology. *Journal of the History of Biology*, *31*(3), 305–325. doi: 10.1023/a:1004312823993

Gelman, R., & Carey, S. (Eds.). (1991). *The Epigenesis of Mind: Essays on Biology and Cognition*. New York: Psychology Press.

Gelman, S. A. (2003). *The Essential Child: Origins of Essentialism in Everyday Thought*. Oxford: Oxford University Press.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Gilbert, S. F. (1994). Dobzhansky, Waddington, and Schmalhausen: Embryology and the Modern Synthesis. In M. B. Adams (Ed.), *The Evolution of Theodosius Dobzhansky: Essays on His Life and Thought in Russia and America* (pp. 143–154). Princeton: Princeton University Press.

Gilbert, S. F. (2001). Ecological Developmental Biology: Developmental Biology Meets the Real World. *Developmental Biology*, *233*(1), 1–12. doi: 10.1006/dbio.2001.0210

Gilbert, S. F. (2003). Evo-Devo, Devo-Evo, and Devgen-Popgen. *Biology & Philosophy*, *18*(2), 347–352. doi: 10.1023/A:1023944304419

Gilbert, S. F. (2012). Ecological developmental biology: Environmental signals for normal animal development. *Evolution & development*, *14*(1), 20–28. doi: 10.1111/j.1525-142X.2011.00519.x

Gilbert, S. F., & Bard, J. B. L. (2014). Formalizing theories of development: A fugue on the orderliness of change. In A. Minelli & T. Pradeu (Eds.), *Towards a Theory of Development* (pp. 129–143). Oxford: Oxford University Press.

Gilbert, S. F., & Barresi, M. J. F. (2010). *Developmental Biology* (11th ed.). Sunderland, MA: Sinauer.

Gilbert, S. F., & Epel, D. (2015). *Ecological Developmental Biology: The Environmental Regulation of Development, Health, and Evolution*. Sunderland, MA: Sinauer.

Gilbert, S. F., Sapp, J., & Tauber, A. I. (2012). A symbiotic view of life: We have never been individuals. *The Quarterly Review of Biology*, *87*(4), 325–341.

doi: 10.1086/668166

Gilbert, S. F., & Sarkar, S. (2000). Embracing complexity: Organicism for the 21st century. *Developmental Dynamics*, *219*(1), 1–9. doi: 10.1002/1097-0177(2000)9999:9999<::AID-DVDY1036>3.0.CO;2-A.

Gilbert, S. F., & Tauber, A. I. (2016). Rethinking individuality: The dialectics of the holobiont. *Biology & Philosophy*, *31*(6), 839–853. doi: 10.1007/s10539-016-9541-3

Gissis, S. B., & Jablonka, E. (Eds.). (2011). *Transformations of Lamarckism: From Subtle Fluids to Molecular Biology*. Cambridge, MA: The MIT Press.

Godfrey-Smith, P. (1994). A modern history theory of functions. *Noûs*, *28*(3), 344–362. doi: 10.2307/2216063

Godfrey-Smith, P. (1996). *Complexity and the Function of Mind in Nature*. Cambridge: Cambridge University Press.

Godfrey-Smith, P. (2009). *Darwinian Populations and Natural Selection*. Oxford: Oxford University Press.

Godfrey-Smith, P. (2014). Sender-receiver systems within and between organisms. *Philosophy of Science*, *81*(5), 866–878. doi: 10.1086/677686

Godfrey-Smith, P. (2016a). Individuality, subjectivity, and minimal cognition. *Biology & Philosophy*, *31*(6), 775–796. doi: 10.1007/s10539-016-9543-1

Godfrey-Smith, P. (2016b). Mind, matter, and metabolism. *The Journal of Philosophy*, *113*(10), 481–506. doi: 10.5840/jphil20161131034

Godfrey-Smith, P. (2017). The subject as cause and effect of evolution. *Interface Focus*, *7*(5), 20170022. doi: 10.1098/rsfs.2017.0022

Godfrey-Smith, P. (2019). Evolving across the explanatory gap. *Philosophy, Theory, and Practice in Biology*, *11*(001). doi: 10.3998/ptpbio.16039257.0011.001

Goldschmidt, R. (1982). *The Material Basis of Evolution*. New Haven, CT: Yale University Press. (Reprinted 1940 edition with an introduction by S. J. Gould.)

Goldstein, K. (1934). *Der Aufbau des Organismus: Einführung in die Biologie unter besonderer Berücksichtigung der Erfahrungen am kranken Menschen*. Den Haag: Nijhoff. (English translation in Goldstein, 1939.)

Goldstein, K. (1939). *The Organism: A Holistic Approach to Biology Derived from Pathological Data in Man*. New York: American Book Company.

Goodwin, B. (1994). *How the Leopard Changed Its Spots: The Evolution of Complexity*. London: Weidenfeld & Nicholson.

Gottlieb, G. (1991). Epigenetic systems view of human development. *Developmental Psychology*, *27*(1), 33–34. doi: 10.1037/0012-1649.27.1.33

Gottlieb, G. (1997). *Synthesizing Nature-Nurture: Prenatal Roots of Instinctive Behavior*. Mahwah, NJ: Lawrence Erlbaum.

Gottlieb, G. (1998). Normally occurring environmental and behavioral influences on gene activity: From Central Dogma to Probabilistic Epigenesis. *Psychological Review*, *105*(4), 792–802. doi: 10.1037/0033-295X.105.4.792-802

Gottlieb, G. (2002). On the epigenetic evolution of species-specific perception: The developmental manifold concept. *Cognitive Development*, *17*(3-4), 1287–1300. doi: 10.1016/s0885-2014(02)00120-x

Gottlieb, G. (2007). Probabilistic epigenesis. *Developmental Science*, *10*(1), 1–11. doi: 10.1111/j.1467-7687.2007.00556.x

Gould, S. J. (1977). *Ontogeny and Phylogeny*. Cambridge, MA: Harvard University Press.

Gould, S. J. (1983). *Hen's Teeth and Horse's Toes: Further Reflections in Natural History*. New York: W. W. Norton & Company.

Gould, S. J. (1989). Punctuated equilibrium in fact and theory. *Journal of Social and Biological Structures*, *12*(2-3), 117–136. doi: 10.1016/0140-1750(89)90040-7

Gould, S. J. (1992). *Ever Since Darwin: Reflections in Natural History*. New York: W. W. Norton & Company.

Gould, S. J. (2002). *The Structure of Evolutionary Theory*. Cambridge, MA: Harvard University Press.

Gould, S. J., & Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, *205*(1161), 581–598. doi: 10.1098/rspb.1979.0086

Gould, S. J., & Vrba, E. S. (1982). Exaptation—a missing term in the science of form. *Paleobiology*, *8*(1), 4–15. doi: 10.1017/s0094837300004310

Gray, R. (1992). Death of the gene: Developmental systems strike back. In P. E. Griffiths (Ed.), *Trees of Life: Essays in Philosophy of Biology* (pp. 165–209). Dordrecht: Kluwer.

Grice, P. (1991). *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Griesemer, J. R. (2000a). Development, culture, and the units of inheritance. *Philosophy of Science*, *67*(S3), S348–S368. doi: 10.1086/392831

Griesemer, J. R. (2000b). The units of evolutionary transition. *Selection*, *1*(1-3), 67–80. doi: 10.1556/select.1.2000.1-3.7

Griesemer, J. R. (2014). Reproduction and scaffolded developmental processes: an integrated evolutionary perspective. In A. Minelli & T. Pradeu (Eds.), *Towards a Theory of Development* (pp. 183–202). Oxford: Oxford University Press.

Griesemer, J. R. (2016). Reproduction in complex life cycles: Toward a developmental reaction norms perspective. *Philosophy of Science*, *83*(5), 803–815.

doi: 10.1086/687865

Griffith, V., & Koch, C. (2014). Quantifying synergistic mutual information. In M. Prokopenko (Ed.), *Guided Self-Organization: Inception* (pp. 159–190). Berlin: Springer.

Griffiths, P. E. (1993). Functional analysis and proper functions. *The British Journal for the Philosophy of Science*, *44*(3), 409–422. doi: 10.1093/bjps/44.3.409

Griffiths, P. E. (2002). What is innateness? *The Monist*, *85*(1), 70–85. doi: 10.5840/monist20028518

Griffiths, P. E. (2013). Lehrman's dictum: information and explanation in developmental biology. *Developmental Psychobiology*, *55*(1), 22–32. doi: 10.1002/dev.21087

Griffiths, P. E. (2016). Proximate and ultimate information in biology. In M. Couch & J. Pfeifer (Eds.), *The Philosophy of Philip Kitcher* (pp. 74–97). Oxford: Oxford University Press.

Griffiths, P. E. (2017). Genetic, epigenetic and exogenetic information in development and evolution. *Interface Focus*, *7*(5), 20160152. doi: 10.1098/rsfs.2016.0152

Griffiths, P. E., & Gray, R. (1994). Developmental systems and evolutionary explanation. *The Journal of Philosophy*, *91*(6), 277–304. doi: 10.2307/2940982

Griffiths, P. E., & Knight, R. (1998). What is the developmentalist challenge? *Philosophy of Science*, *65*(2), 253–258. doi: 10.1086/392636

Griffiths, P. E., Pocheville, A., Calcott, B., Stotz, K., Kim, H., & Knight, R. (2015). Measuring causal specificity. *Philosophy of Science*, *82*(4), 529–555. doi: 10.1086/682914

Griffiths, P. E., & Stotz, K. (2006). Genes in the postgenomic era. *Theoretical Medicine and Bioethics*, *27*(6), 499–521. doi: 10.1007/s11017-006-9020-y

Griffiths, P. E., & Stotz, K. (2013). *Genetics and Philosophy. An Introduction.* Cambridge: Cambridge University Press.

Gupta, M., Prasad, N. G., Dey, S., Joshi, A., & Vidya, T. N. C. (2017). Niche construction in evolutionary theory: The construction of an academic niche? *Journal of Genetics*, *96*(3), 491–504. doi: 10.1007/s12041-017-0787-6

Gánti, T. (2003). *The Principles of Life.* Oxford: Oxford University Press.

Haeckel, E. (1866). *Generelle Morphologie der Organismen. Allgemeine Grundzüge der organischen Formen-Wissenschaft, mechanisch begründet durch die von Charles Darwin reformirte Descendenz-Theorie* (Vol. 1). Berlin: Georg Reimer.

Haldane, J. S. (1931). *The Philosophical Basis of Biology.* London: Hodder & Stoughton.

Hall, B. K. (2000). Guest editorial: Evo-devo or devo-evo — does it matter? *Evolution & Development*, *2*(4), 177–178. doi: 10.1046/j.1525-142x.2000.00003e.x

Hamburger, V. (1980). Embryology and the Modern Synthesis in evolutionary theory. In E. Mayr & W. B. Provine (Eds.), *The Evolutionary Synthesis: Perspectives on the Unification of Biology* (pp. 97–112). Cambridge, MA: Harvard University Press. doi: 10.1007/978-1-4899-6743-5_18

Haugeland, J. (1981). Semantic engines: An introduction to mind design. In J. Haugeland (Ed.), *Mind Design: Philosophy, Psychology, Artificial Intelligence* (pp. 1–34). Cambridge, MA: The MIT Press.

Heras-Escribano, M. (2019). *The Philosophy of Affordances*. Cham: Palgrave Macmillan.

Heras-Escribano, M., & De Jesus, P. (2018). Biosemiotics, the extended synthesis, and ecological information: Making sense of the organism-environment relation at the cognitive level. *Biosemiotics*, *11*(2), 245–262. doi: 10.1007/s12304-018-9322-2

Heras-Escribano, M., & de Pinedo-García, M. (2018). Affordances and landscapes: Overcoming the nature-culture dichotomy through niche construction theory. *Frontiers in Psychology*, *8*, 2294. doi: 10.3389/fpsyg.2017.02294

Hochman, A. (2012). The phylogeny fallacy and the ontogeny fallacy. *Biology & Philosophy*, *28*(4), 593–612. doi: 10.1007/s10539-012-9325-3

Hoffmeyer, J. (1998). Surfaces inside surfaces. On the origin of agency and life. *Cybernetics & Human Knowing*, *5*(1), 33–42.

Hoffmeyer, J. (2002). The central dogma: A joke that became real. *Semiotica*, *138*, 1–13. doi: 10.1515/semi.2002.004

Hoffmeyer, J. (2008a). *Biosemiotics: An Examination into the Signs of Life and the Life of Signs*. Scranton, PA: University of Scranton Press.

Hoffmeyer, J. (2008b). Semiotic scaffolding of living systems. In M. Barbieri (Ed.), *Introduction to Biosemiotics: The New Biological Synthesis* (pp. 149–166). Dordrecht: Springer.

Hoffmeyer, J. (2011a). Biology is immature biosemiotics. In C. Emmeche & K. Kull (Eds.), *Towards a Semiotic Biology: Life is the Action of Signs* (pp. 43–65). London: Imperial College Press.

Hoffmeyer, J. (2011b). The natural history of intentionality. A biosemiotic approach. In T. Schilhab, F. Stjernfelt, & T. W. Deacon (Eds.), *The Symbolic Species Evolved* (pp. 97–116). Dordrecht: Springer.

Huneman, P. (2010). Assessing the prospects for a return of organisms in evolutionary biology. *History and Philosophy of the Life Sciences*, *32*(2-3), 341–371. doi: 10.2307/23335078

Huneman, P. (2017). Why would we call for a new evolutionary synthesis?

The variation issue and the explanatory alternatives. In P. Huneman & D. M. Walsh (Eds.), *Challenging the Modern Synthesis. Adaptation, Development, and Inheritance* (pp. 68–110). Oxford: Oxford University Press. doi: 10.1093/oso/9780199377176.003.0002

Huneman, P., & Walsh, D. M. (Eds.). (2017). *Challenging the Modern Synthesis. Adaptation, Development, and Inheritance*. Oxford: Oxford University Press.

Huxley, J. (1942). *Evolution: The Modern Synthesis*. London: George Allen & Unwin.

Jablonka, E. (2002). Information: Its interpretation, its inheritance, and its sharing. *Philosophy of Science*, *69*(4), 578–605. doi: 10.1086/344621

Jablonka, E. (2004). From replicators to heritably varying phenotypic traits: The extended phenotype revisited. *Biology & Philosophy*, *19*(3), 353–375. doi: 10.1023/B:BIPH.0000036112.02199.7b

Jablonka, E. (2007). The developmental construction of heredity. *Developmental Psychobiology*, *49*(8), 808–817. doi: 10.1002/dev.20260

Jablonka, E., & Lamb, M. J. (1995). *Epigenetic Inheritance and Evolution: The Lamarckian Dimension*. Oxford: Oxford University Press.

Jablonka, E., & Lamb, M. J. (2002). The changing concept of epigenetics. *Annals of the New York Academy of Sciences*, *981*(1), 82–96. doi: 10.1111/j.1749-6632.2002.tb04913.x

Jablonka, E., & Lamb, M. J. (2014). *Evolution in Four Dimensions, Revised Edition: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*. Cambridge, MA: The MIT Press.

Jablonka, E., & Lamb, M. J. (2020). *Inheritance Systems and the Extended Evolutionary Synthesis*. Cambridge: Cambridge University Press.

Jablonka, E., & Raz, G. (2009). Transgenerational epigenetic inheritance: Prevalence, mechanisms, and implications for the study of heredity and evolution. *The Quarterly Review of Biology*, *84*(2), 131–176. doi: 10.1086/598822

Jacob, F. (1993). *The Logic of Life: A History of Heredity*. Princeton, NJ: Princeton University Press.

Johannsen, W. (1909). *Elemente der exakten Erblichkeitslehre*. Jena: Gustav Fischer.

Johansson, L.-G. (2007). Causation: A synthesis of three approaches. In *Computation, Information, Cognition: The Nexus and the Liminal* (pp. 75–86). Newcastle upon Tyne: Cambridge Scholars Publishing.

Johnston, T. D. (1987). The persistence of dichotomies in the study of behavioral development. *Developmental Review*, *7*(2), 149–182. doi: 10.1016/0273-2297(87)90011-6

Johnston, T. D. (2010). Developmental systems theory. In M. S. Blumberg,

J. H. Freeman, & S. R. Robinson (Eds.), *Oxford Handbook of Developmental Behavioral Neuroscience* (pp. 12–29). Oxford: Oxford University Press.

Johnston, T. D., & Edwards, L. (2002). Genes, interactions, and the development of behavior. *Psychological Review*, *109*(1), 26–34. doi: 10.1037/0033-295x .109.1.26

Juarrero, A. (1999). *Dynamics in Action: Intentional Behavior as a Complex System*. Cambridge, MA: The MIT Press.

Juarrero, A. (2009). Top-down causation and autonomy in complex systems. In N. Murphy, G. F. R. Ellis, & T. O'Connor (Eds.), *Downward Causation and the Neurobiology of Free Will* (pp. 83–102). Berlin: Springer.

Kant, I. (2007). *Critique of Judgement*. Oxford: Oxford University Press.

Karmiloff-Smith, A. (2006). The tortuous route from genes to behavior: A neuroconstructivist approach. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(1), 9–17. doi: 10.3758/cabn.6.1.9

Karmiloff-Smith, A. (2009). Nativism versus neuroconstructivism: Rethinking the study of developmental disorders. *Developmental Psychology*, *45*(1), 56–63. doi: 10.1037/a0014506

Kauffman, S. (1993). *The Origins of Order: Self-Organization and Selection in Evolution*. New York: Oxford University Press.

Kauffman, S. (1995). *At Home in the Universe: The Search for Laws of Self-Organization and Complexity*. New York: Oxford University Press.

Kauffman, S. (2000). *Investigations*. Oxford: Oxford University Press.

Kauffman, S. (2003). Molecular autonomous agents. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *361*(1807), 1089–1099. doi: 10.1098/rsta.2003.1186

Kauffman, S. (2019). *A World Beyond Physics: The Emergence and Evolution of Life*. Oxford: Oxford University Press.

Keller, E. F. (2002). *The Century of the Gene*. Cambridge, MA: Harvard University Press.

Keller, E. F. (2010). *The Mirage of a Space between Nature and Nurture*. Durham, NC: Duke University Press.

Keller, E. F. (2014). From gene action to reactive genomes. *The Journal of Physiology*, *592*(11), 2423–2429. doi: 10.1113/jphysiol.2014.270991

Kendal, J. R. (2011). Cultural niche construction and human learning environments: Investigating sociocultural perspectives. *Biological Theory*, *6*(3), 241–250. doi: 10.1007/s13752-012-0038-2

Kerr, B. (2007). Niche construction and cognitive evolution. *Biological Theory*, *2*(3), 250–262. doi: 10.1162/biot.2007.2.3.250

Kirschner, M. W., & Gerhart, J. C. (2008). *The Plausibility of Life: Resolving Darwin's Dilemma*. New Haven: Yale University Press.

Kitano, H. (Ed.). (2001). *Foundations of Systems Biology.* Cambridge, MA: The MIT Press.

Klopfer, P. H. (2001). Parental care and development. In S. Oyama, P. E. Griffiths, & R. D. Gray (Eds.), *Cycles of Contingency: Developmental Systems and Evolution* (pp. 167–173). Cambridge, MA: The MIT Press.

Korbak, T. (2015). Scaffolded minds and the evolution of content in signaling pathways. *Studies in Logic, Grammar and Rhetoric*, *41*(1), 89–103. doi: 10.1515/slgr-2015-0022

Kováč, L. (2006). Life, chemistry and cognition: Conceiving life as knowledge embodied in sentient chemical systems might provide new insights into the nature of cognition. *EMBO reports*, *7*(6), 562–566. doi: 10.1038/sj.embor.7400717

Kripke, S. (1980). *Naming and Necessity.* Cambridge, MA: Harvard University Press.

Krohs, U., & Kroes, P. (Eds.). (2009). *Functions in Biological and Artificial Worlds: Comparative Philosophical Perspectives.* Cambridge, MA: The MIT Press.

Kull, K. (1999). Biosemiotics in the twentieth century: A view from biology. *Semiotica*, *127*(1-4), 385–414. doi: 10.1515/semi.1999.127.1-4.385

Kull, K. (2001). Jakob von Uexküll: An introduction. *Semiotica*, *2001*(134), 1–59. doi: 10.1515/semi.2001.013

Kull, K. (2009). Biosemiotics: To know, what life knows. *Cybernetics & Human Knowing*, *16*(3-4), 81–88.

Kull, K. (2011). Life is many, and sign is essentially plural: On the methodology of biosemiotics. In C. Emmeche & K. Kull (Eds.), *Towards a Semiotic Biology: Life Is the Action of Signs* (pp. 113–129). London: Imperial Colleague Press.

Kull, K. (2021). Natural selection and self-organization do not make meaning, while the agent's choice does. *Biosemiotics*, *14*(1), 49–53. doi: 10.1007/s12304-021-09422-2

Kull, K., Emmeche, C., & Favareau, D. (2011). Biosemiotic research questions. In C. Emmeche & K. Kull (Eds.), *Towards a Semiotic Biology: Life Is the Action of Signs* (pp. 67–90). London: Imperial College Press. doi: 10.1142/9781848166882_0004

Kuo, Z. Y. (1921). Giving up instincts in psychology. *The Journal of Philosophy*, *18*(24), 645–664. doi: 10.2307/2939656

Kuo, Z. Y. (1922). How are our instincts acquired? *Psychological Review*, *29*(5), 344–365. doi: 10.1037/h0073689

Kuo, Z. Y. (1932). Ontogeny of embryonic behavior in Aves: I. The chronology and general nature of the behavior of the chick embryo. *The Journal of Experimental Zoology*, *61*(3), 395–430. doi: 10.1002/jez.1400610304

Kuo, Z. Y. (1976). *The Dynamics of Behavior Development: An Epigenetic View* (2nd ed.). New York: Plenum Press.

Ladyman, J., & Wiesner, K. (2020). *What Is a Complex System?* New Haven, CT: Yale University Press.

Laland, K. N., Matthews, B., & Feldman, M. W. (2016). An introduction to Niche Construction Theory. *Evolutionary Ecology*, *30*(2), 191–202. doi: 10.1007/s10682-016-9821-z

Laland, K. N., Odling-Smee, J., & Endler, J. A. (2017). Niche construction, sources of selection and trait coevolution. *Interface Focus*, *7*(5), 20160147. doi: 10.1098/rsfs.2016.0147

Laland, K. N., Odling-Smee, J., & Feldman, M. W. (2019). Understanding niche construction as an evolutionary process. In T. Uller & K. N. Laland (Eds.), *Evolutionary Causation: Biological and Philosophical Reflections* (pp. 127–152). Cambridge, MA: The MIT Press.

Laland, K. N., Odling-Smee, J., Hoppitt, W., & Uller, T. (2013). More on how and why: A response to commentaries. *Biology & Philosophy*, *28*(5), 793–810. doi: 10.1007/s10539-013-9380-4

Laland, K. N., & Sterelny, K. (2007). Seven reasons (not) to neglect niche construction. *Evolution*, *60*(9), 1751–1762. doi: 10.1111/j.0014-3820.2006.tb00520.x

Laland, K. N., Uller, T., Feldman, M., Sterelny, K., Müller, G. B., Moczek, A., … Strassman, J. E. (2014). Does evolutionary theory need a rethink? *Nature*, *514*(7521), 161–164. doi: 10.1038/514161a

Laubichler, M. D., & Maienschein, J. (Eds.). (2007). *From Embryology to Evo-Devo: A History of Developmental Evolution.* Cambridge, MA: The MIT Press.

Lehrman, D. S. (1953). A critique of Konrad Lorenz's theory of instinctive behavior. *The Quarterly Review of Biology*, *28*(4), 337–363. doi: 10.1086/399858

Lehrman, D. S. (1970). Semantic and conceptual issues in the nature-nurture problem. In L. R. Aronson, E. Tobach, D. S. Lehrman, & J. Rosenblatt (Eds.), *Development and Evolution of Behavior. Essays in Memory of T. C. Schneirla* (pp. 17–52). San Francisco: W. H. Freeman and Company.

Lennox, J. G. (1993). Darwin was a teleologist. *Biology & Philosophy*, *8*(4), 409–421. doi: 10.1007/BF00857687

Lenoir, T. (1980). Kant, Blumenbach, and vital materialism in German biology. *Isis*, *71*(1), 77–108. doi: 10.1086/352408

Lenoir, T. (1981). Teleology without regrets. The transformation of physiology in Germany: 1790–1847. *Studies in History and Philosophy of Science Part A*, *12*(4), 293–354. doi: 10.1016/0039-3681(81)90019-4

Lenoir, T. (1989). *The Strategy of Life: Teleology and Mechanics in Nineteenth-Century German Biology.* Chicago: The University of Chicago Press.

Levin, M., & Dennett, D. C. (2020). Cognition all the way down. *Aeon*. Retrieved from `https://aeon.co/essays/how-to-understand-cells-tissues-and-organisms-as-agents-with-agendas`

Levin, M., Keijzer, F., Lyon, P., & Arendt, D. (2021). *Uncovering cognitive similarities and differences, conservation and innovation* (Vol. 376) (No. 1821). The Royal Society. doi: 10.1098/rstb.2020.0458

Levins, R., & Lewontin, R. C. (1985). *The Dialectical Biologist*. Cambridge, MA: Harvard University Press.

Lewis, D. (1969). *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.

Lewontin, R. C. (1970). The units of selection. *Annual Review of Ecology and Systematics*, *1*, 1–18. doi: 10.1146/annurev.es.01.110170.000245

Lewontin, R. C. (1974a). The analysis of variance and the analysis of causes. *American Journal of Human Genetics*, *26*(3), 400–411.

Lewontin, R. C. (1974b). *The Genetic Basis of Evolutionary Change*. New York: Columbia University Press.

Lewontin, R. C. (1978). Adaptation. *Scientific American*, *239*(3), 212–231. doi: 10.1038/scientificamerican0978-212

Lewontin, R. C. (1983a). Gene, organism, and environment. In D. S. Bendall (Ed.), *Evolution from Molecules to Men* (pp. 273–285). Cambridge: Cambridge University Press. (Reprinted in Oyama et al., 2001, 59–66)

Lewontin, R. C. (1983b). The organism as the subject and object of evolution. *Scientia, rivista internazionale di sintesi scientifica*, *118*, 65–95. (Reprinted in Levins & Lewontin, 1985, pp. 85–106)

Lewontin, R. C. (2000). *The Triple Helix: Gene, Organism, and Environment*. Cambridge, MA: Harvard University Press.

Lewontin, R. C. (2010, May 27). Not so natural selection. *New York Review of Books*.

Lewontin, R. C., & Levins, R. (2007). *Biology under the Influence: Dialectical Essays on Ecology, Agriculture, and Health*. New York: Monthly Review Press.

Lickliter, R., & Berry, T. D. (1990). The phylogeny fallacy: Developmental psychology's misapplication of evolutionary theory. *Developmental Review*, *10*(4), 348–364. doi: 10.1016/0273-2297(90)90019-Z

Linquist, S. (2018). The conceptual critique of innateness. *Philosophy Compass*, *13*(5), e12492. doi: 10.1111/phc3.12492

Lloyd, E. A. (2021). *Adaptation*. Cambridge: Cambridge University Press.

Lloyd, E. A., & Gould, S. J. (2017). Exaptation revisited: Changes imposed by evolutionary psychologists and behavioral biologists. *Biological Theory*, *12*(1), 50–65. doi: 10.1007/s13752-016-0258-y

Lorenzo, G., & Longa, V. M. (2018). *El innatismo. Origen, variaciones y vitalidad de una idea*. Madrid: Cátedra.

Lotfi, S. (2010). The 'purposiveness' of life: Kant's critique of natural teleology. *The Monist*, *93*(1), 123–134. doi: 10.5840/monist20109318

Love, A. C. (2007). Functional homology and homology of function: Biological concepts and philosophical consequences. *Biology & Philosophy*, *22*(5), 691–708. doi: 10.1007/s10539-007-9093-7

Love, A. C. (2008). Explaining evolutionary innovations and novelties: Criteria of explanatory adequacy and epistemological prerequisites. *Philosophy of Science*, *75*(5), 874–886. doi: 10.1086/594531

Love, A. C. (Ed.). (2015). *Conceptual Change in Biology: Scientific and Philosophical Perspectives on Evolution and Development*. Dordrecht: Springer.

Lumsden, C. J., & Wilson, E. O. (1981). *Genes, Mind, and Culture. The Coevolutionary Process*. Cambridge, MA: Harvard University Press.

Lyon, P. (2015). The cognitive cell: Bacterial behavior reconsidered. *Frontiers in Microbiology*, *6*, 264. doi: 10.3389/fmicb.2015.00264

Macdonald, G., & Papineau, D. (2006). Introduction: Prospects and problems for teleosemantics. In G. Macdonald & D. Papineau (Eds.), *Teleosemantics. New Philosophical Essays* (p. 1-22). Oxford: Oxford University Press.

Mameli, M. (2005). The inheritance of features. *Biology & Philosophy*, *20*(2-3), 365–399. doi: 10.1007/s10539-004-0560-0

Mameli, M. (2007). Genes, environments, and concepts of biological inheritance. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The Innate Mind* (Vol. 3: *Foundations and the Future*, pp. 37–54). Oxford: Oxford University Press.

Mameli, M. (2008). On innateness: The clutter hypothesis and the cluster hypothesis. *The Journal of Philosophy*, *105*(12), 719–736. doi: 10.5840/jphil20081051216

Mameli, M., & Bateson, P. (2006). Innateness and the sciences. *Biology & Philosophy*, *21*(2), 155–188. doi: 10.1007/s10539-005-5144-0

Mameli, M., & Bateson, P. (2011). An evaluation of the concept of innateness. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1563), 436–443. doi: 10.1098/rstb.2010.0174

Marcus, G. (2004). *The Birth of the Mind: How a Tiny Number of Genes Creates The Complexities of Human Thought*. New York: Basic Books.

Margolis, E. (1998). How to acquire a concept. *Mind & Language*, *13*(3), 347–369. doi: 10.1111/1468-0017.00081

Margolis, E., & Laurence, S. (2011). Learning matters: The role of learning in concept acquisition. *Mind & Language*, *26*(5), 507–539. doi: 10.1111/j.1468-0017.2011.01429.x

Margulis, L. (1990). Words as battle cries—symbiogenesis and the new field of

endocytobiology. *Bioscience*, *40*(9), 673–677. doi: 10.2307/1311435

Margulis, L. (1993). *Symbiosis in Cell Evolution: Microbial Communities in the Archean and Proterozoic Eons*. San Francisco: W. H. Freeman and Company.

Markoš, A. (2002). *Readers of the Book of Life: Contextualizing Developmental Evolutionary Biology*. New York: Oxford University Press.

Martínez, M. (2013a). Teleosemantics and Indeterminacy. *Dialectica*, *67*(4), 427–453. doi: 10.1111/1746-8361.12039

Martínez, M. (2013b). Teleosemantics and productivity. *Philosophical Psychology*, *26*(1), 47–68. doi: 10.1080/09515089.2011.625115

Martínez, M. (2015). Deception in sender-receiver games. *Erkenntnis*, *80*(1), 215–227. doi: 10.1007/s10670-014-9623-z

Matthen, M., & Ariew, A. (2002). Two ways of thinking about fitness and natural selection. *The Journal of Philosophy*, *99*(2), 55–83. doi: 10.2307/3655552

Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and Cognition. The Realization of the Living*. Dordrecht: Reidel.

Maynard Smith, J. (1969). The status of neo-Darwinism. In C. H. Waddington (Ed.), *Sketching Theoretical Biology: Towards a Theoretical Biology* (Vol. 2, pp. 82–89). New Brunswick, NJ: Transaction.

Maynard Smith, J. (1978). Optimization theory in evolution. *Annual Review of Ecology and Systematics*, *9*, 31–56. doi: 10.1146/annurev.es.09.110178.000335

Maynard Smith, J. (2000). The concept of information in biology. *Philosophy of Science*, *67*(2), 177–194. doi: 10.1086/392768

Maynard Smith, J., Burian, R., Kauffman, S., Alberch, P., Campbell, J. H., Goodwin, B., … Wolpert, L. (1985). Developmental constraints and evolution: A perspective from the Mountain Lake Conference on Development and Evolution. *The Quarterly Review of Biology*, *60*(3), 265–287. doi: 10.1086/414425

Mayr, E. (1961). Cause and effect in biology: Kinds of causes, predictability, and teleology are viewed by a practicing biologist. *Science*, *134*(3489), 1501–1506. doi: 10.1126/science.134.3489.1501

Mayr, E. (1974). Teleological and teleonomic, a new analysis. In R. S. Cohen & M. W. Wartofsky (Eds.), *Methodological and Historical Essays in the Natural and Social Sciences* (pp. 91–117). Dordrecht: Reidel.

Mayr, E. (1975). Typological versus population thinking. In *Evolution and the Diversity of Life. Selected Essays* (pp. 26–29). Cambridge, MA: Harvard University Press. (Reprinted in Sober, 2006, 325–328.)

Mayr, E. (1980). Prologue: Some thoughts on the history of the Evolutionary Synthesis. In E. Mayr & W. B. Provine (Eds.), *The Evolutionary Synthesis: Perspectives on the Unification of Biology* (pp. 1–48). Cambridge, MA:

Harvard University Press.

Mayr, E. (1988). *Toward a New Philosophy of Biology: Observations of an Evolutionist*. Cambridge, MA: Harvard University Press.

Mayr, E. (1998). The multiple meanings of 'teleological'. *History and Philosophy of the Life Sciences*, *20*(1), 35–40.

Mayr, E. (1999). *Systematics and the Origin of Species from the Viewpoint of a Zoologist* (2nd ed.). Cambridge, MA: Harvard University Press.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, *5*(4), 115–133. doi: 10.1007/BF02478259

McGinn, C. (2015). *Philosophy of Language: The Classics Explained*. Cambridge, MA: The MIT Press.

McLaughlin, P. (1990). *Kant's Critique of Teleology in Biological Explanation: Antinomy and Teleology*. Lewiston, NY: Edwin Mellen Press.

McLaughlin, P. (2000). *What Functions Explain: Functional Explanation and Self-Reproducing Systems*. Cambridge: Cambridge University Press.

McShea, D. W. (2012). Upper-directed systems: A new approach to teleology in biology. *Biology & Philosophy*, *27*(5), 663–684. doi: 10.1007/s10539-012-9326-2

Mendel, G. (1866). Versuche über Pflanzen-Hybriden. *Verhandlungen des naturforschenden Vereines in Brünn*, *IV*(für das Jahr 1865 Abhandlungen), 3–47.

Mercier, H., & Sperber, D. (2017). *The Enigma of Reason*. Cambridge, MA: Harvard University Press.

Meyer-Abich, A. (1950). *Beiträge zur Theorie der Evolution der Organismen: II. Typensynthese durch Holobiose*. Leiden: E. L. Brill.

Michel, G. F. (2013). The role of developmental psychobiology in the unification of psychology. *Review of General Psychology*, *17*(2), 210–215. doi: 10.1037/a0032936

Michel, G. F., & Moore, C. L. (1995). *Developmental Psychobiology: An Interdisciplinary Science*. Cambridge, MA: MIT Press.

Millikan, R. G. (1984). *Language, Thought, and Other Biological Categories. New Foundations for Realism*. Cambridge, MA: The MIT press.

Millikan, R. G. (1989). In defense of proper functions. *Philosophy of Science*, *56*(2), 288–302. doi: 10.1086/289488

Millikan, R. G. (1993). *White Queen Psychology and Other Essays for Alice*. Cambridge, MA: The MIT Press.

Millikan, R. G. (2000a). Naturalizing intentionality. In B. Elevitch (Ed.), *The Proceedings of the Twentieth World Congress of Philosophy* (Vol. IX: Philosophy of Mind, pp. 83–90). Bowling Green: Philosophy Documentation Center. Bowling Green State University. doi: 10.5840/wcp202000997

Millikan, R. G. (2000b). *On Clear and Confused Ideas: An Essay about Substance Concepts*. Cambridge: Cambridge University Press.

Millikan, R. G. (2002). Biofunctions: Two paradigms. In A. Ariew, R. Cummins, & M. Perlman (Eds.), *Functions: New Essays in the Philosophy of Psychology and Biology* (pp. 113–143). Oxford: Oxford University Press.

Millikan, R. G. (2003). Teleological theories of mental content. In L. Nadel (Ed.), *Encyclopedia of Cognitive Science* (pp. 1138–1141). New York: Wiley.

Millikan, R. G. (2004). *Varieties of Meaning*. Cambridge, MA: The MIT press.

Millikan, R. G. (2006). Useless content. In G. Macdonald & D. Papineau (Eds.), *Teleosemantics. New Philosophical Essays* (pp. 100–114). Oxford: Oxford University Press.

Millikan, R. G. (2017). *Beyond Concepts: Unicepts, Language, and Natural Information*. Oxford: Oxford University Press.

Millstein, R. L. (2006). Natural selection as a population-level causal process. *The British Journal for the Philosophy of Science*, *57*(4), 627–653. doi: 10.1093/bjps/axl025

Minelli, A. (2009). *The Development Of Animal Form: Ontogeny, Morphology, and Evolution*. Cambridge: Cambridge University Press.

Minelli, A. (2011). Animal Development, an Open-Ended Segment of Life. *Biological Theory*, *6*(1), 4–15. doi: 10.1007/s13752-011-0002-6

Minelli, A., & Fusco, G. (Eds.). (2008). *Evolving Pathways: Key Themes in Evolutionary Developmental Biology*. Cambridge: Cambridge University Press.

Mivart, S. G. (1871). *On the Genesis of Species*. London: Macmillan.

Moczek, A. P. (2008). On the origins of novelty in development and evolution. *BioEssays*, *30*(5), 432–447. doi: 10.1002/bies.20754

Moczek, A. P. (2009). Phenotypic plasticity and the origins of diversity: A case study of horned beetles. In D. W. Whitman & T. N. Ananthakrishnan (Eds.), *Phenotypic Plasticity of Insects: Mechanisms and Consequences* (pp. 81–134). Enfield, NH: Science Publishers.

Moczek, A. P., Sultan, S. E., Foster, S., Ledón-Rettig, C., Dworkin, I., Nijhout, H. F., … Pfennig, D. W. (2011). The role of developmental plasticity in evolutionary innovation. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1719), 2705–2713. doi: 10.1098/rspb.2011.0971

Monod, J. (1971). *Chance and Necessity: Essay on the Natural Philosophy of Modern Biology*. New York: Knopf.

Monod, J., & Jacob, F. (1961). General conclusions: Teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harbor Symposia on Quantitative Biology*, *26*, 389–401. doi: 10.1101/SQB.1961.026.01.048

Moore, C. L. (2003). Evolution, development, and the individual acquisition of

traits: What we've learned. In B. H. Weber & D. J. Depew (Eds.), *Evolution and Learning: The Baldwin Effect Reconsidered* (pp. 115–139). Cambridge, MA: The MIT Press.

Moore, D. S. (2001). *The Dependent Gene: The Fallacy of "Nature vs. Nurture"*. New York: Henry Holt.

Moran, N. A. (1994). Adaptation and constraint in the complex life cycles of animals. *Annual Review of Ecology and Systematics*, *25*, 573–600. doi: 10.1146/annurev.es.25.110194.003041

Moreno, A. (1998). Information, causality and self-reference in natural and artificial systems. *AIP Conference Proceedings*, *437*(1), 202–206. doi: 10.1063/1.56301

Moreno, A., & Mossio, M. (2015). *Biological Autonomy. A Philosophical and Theoretical Inquiry*. Dordrecht: Springer.

Moreno, A., & Ruiz-Mirazo, K. (2011). The informational nature of biological causality. In G. Terzis & R. Arp (Eds.), *Information and Living Systems: Philosophical and Scientific Perspectives* (pp. 157–176). Cambridge, MA: The MIT Press. doi: 10.7551/mitpress/9780262201742.003.0007

Morgan, T. H., Sturtevant, A. H., Muller, H. J., & Bridges, C. B. (1915). *The Mechanism of Mendelian Heredity*. New York: Henry Holt and Company.

Moss, L. (2003). *What Genes Can't Do*. Cambridge, MA: The MIT Press.

Mossio, M., & Bich, L. (2017). What makes biological organisation teleological? *Synthese*, *194*(4), 1089–1114. doi: 10.1007/s11229-014-0594-z

Mossio, M., Montévil, M., & Longo, G. (2016). Theoretical principles for biology: Organization. *Progress in Biophysics and Molecular Biology*, *122*(1), 24–35. doi: 10.1016/j.pbiomolbio.2016.07.005

Mossio, M., Saborido, C., & Moreno, A. (2009). An organizational account of biological functions. *The British Journal for the Philosophy of Science*, *60*(4), 813–841. doi: 10.1093/bjps/axp036

Müller, G. B. (2017). Why an extended evolutionary synthesis is necessary. *Interface Focus*, *7*(5), 20170015. doi: 10.1098/rsfs.2017.0015

Müller, G. B., & Newman, S. A. (Eds.). (2003). *Origination of Organismal Form: Beyond the Gene in Developmental and Evolutionary Biology*. Cambridge, MA: The MIT Press.

Nagel, E. (1979). *Teleology Revisited and Other Essays in the Philosophy and History of Science*. New York: Columbia University Press.

Nanay, B. (2014). Teleosemantics without etiology. *Philosophy of Science*, *81*(5), 798–810. doi: 10.1086/677684

Nanney, D. L. (1958). Epigenetic control systems. *Proceedings of the National Academy of Sciences*, *44*(7), 712–717. doi: 10.1073/pnas.44.7.712

Neander, K. (1991a). Functions as selected effects: The conceptual analyst's

defense. *Philosophy of Science*, *58*(2), 168–184. doi: 10.1086/289610

Neander, K. (1991b). The teleological notion of 'function'. *Australasian Journal of Philosophy*, *69*(4), 454–468. doi: 10.1080/00048409112344881

Neander, K. (1995a). Explaining complex adaptations: A reply to Sober's 'Reply to Neander'. *The British journal for the philosophy of science*, *46*(4), 583–587. doi: 10.1093/bjps/46.4.583

Neander, K. (1995b). Misrepresenting and malfunctioning. *Philosophical Studies*, *79*(2), 109–141. doi: 10.1007/bf00989706

Neander, K. (2006a). Content for cognitive science. In G. Macdonald & D. Papineau (Eds.), *Teleosemantics. New Philosophical Essays* (pp. 167–194). Oxford: Oxford University Press.

Neander, K. (2006b). Naturalistic theories of reference. In M. Devitt & R. Hanley (Eds.), *The Blackwell Guide to the Philosophy of Language* (pp. 374–391). Malden, MA: Wiley-Blackwell.

Neander, K. (2007). Biological approaches to mental representation. In M. Matthen & C. Stephens (Eds.), *Handbook of the Philosophy of Science* (Vol. 3: Philosophy of Biology, pp. 549–565). Amsterdam: Elsevier.

Neander, K. (2008). Teleological theories of mental content: Can Darwin solve the problem of intentionality? In M. Ruse (Ed.), *The Oxford Handbook of Philosophy of Biology* (pp. 381–409). Oxford.

Neander, K. (2013). Toward an informational teleosemantics. In D. Ryder, J. Kingsbury, & K. Williford (Eds.), *Millikan and Her Critics* (pp. 21–40). Malden, MA: Wiley-Blackwell.

Neander, K. (2017a). Functional analysis and the species design. *Synthese*, *194*(4), 1147–1168. doi: 10.1007/s11229-015-0940-9

Neander, K. (2017b). *A Mark of the Mental. In Defense of Informational Teleosemantics*. Cambridge, MA: The MIT Press.

Neander, K. (2018). Does biology need teleology? In R. Joyce (Ed.), *The Routledge Handbook of Evolution and Philosophy* (pp. 64–76). London: Routledge.

Neander, K., & Schulte, P. (2021). Teleological theories of mental content. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2021 ed.). Metaphysics Research Lab, Stanford University. `https://plato.stanford.edu/archives/spr2021/entries/content-teleological/`.

Needham, J. (1936). *Order and Life*. Cambridge: Cambridge University Press.

Newman, S. A. (2012). Physico-genetic determinants in the evolution of development. *Science*, *338*(6104), 217–219. doi: 10.1126/science.1222003

Newman, S. A. (2021). Inherency. In L. Nuño de la Rosa & G. B. Müller (Eds.), *Evolutionary Developmental Biology: A Reference Guide* (pp. 121–132). Cham: Springer. doi: 10.1007/978-3-319-32979-6_78

Newman, S. A. (2022a). Inherency and agency in the origin and evolution of

biological functions. *Biological Journal of the Linnean Society*, *blac109*. doi: 10.1093/biolinnean/blac109

Newman, S. A. (2022b). Self-organization in embryonic development: Myth and reality. In A. D. Malassé (Ed.), *Self-Organization as a New Paradigm in Evolutionary Biology: From Theory to Applied Cases in the Tree of Life* (pp. 195–222). Cham: Springer.

Newman, S. A., & Comper, W. D. (1990). 'Generic' physical mechanisms of morphogenesis and pattern formation. *Development*, *110*(1), 1–18. doi: 10.1242/dev.110.1.1

Newman, S. A., Forgacs, G., & Müller, G. B. (2003). Before programs: The physical origination of multicellular forms. *International Journal of Developmental Biology*, *50*(2-3), 289–299. doi: 10.1387/ijdb.052049sn

Newman, S. A., Glimm, T., & Bhat, R. (2018). The vertebrate limb: An evolving complex of self-organizing systems. *Progress in Biophysics and Molecular Biology*, *137*, 12–24. doi: 10.1016/j.pbiomolbio.2018.01.002

Nicholson, D. J. (2012). The concept of mechanism in biology. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *43*(1), 152–163. doi: 10.1016/j.shpsc.2011.05.014

Nicholson, D. J. (2013). Organisms $\neq$ machines. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *44*(4), 669–678. doi: 10.1016/j.shpsc.2013.05.014

Nicholson, D. J. (2014). The return of the organism as a fundamental explanatory concept in biology. *Philosophy Compass*, *9*(5), 347–359. doi: 10.1111/phc3.12128

Nicholson, D. J. (2018). Reconceptualizing the organism: From complex machine to flowing stream. In D. J. Nicholson & J. Dupré (Eds.), *Everything Flows: Towards a Processual Philosophy of Biology* (pp. 138–166). Oxford: Oxford University Press.

Nicholson, D. J., & Gawne, R. (2015). Neither logical empiricism nor vitalism, but organicism: What the philosophy of biology was. *History and Philosophy of the Life Sciences*, *37*(4), 345–381. doi: 10.1007/s40656-015-0085-7

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

Nijhout, H. F. (2003). Development and evolution of adaptive polyphenisms. *Evolution and Development*, *5*(1), 9–18. doi: 10.1046/j.1525-142x.2003.03003.x

Noble, D. (2006). *The Music of Life: Biology Beyond Genes*. Oxford: Oxford University Press.

Noble, D. (2016). *Dance to the Tune of Life: Biological Relativity*. Cambridge:

Cambridge University Press.

Noble, D. (2021). The illusions of the Modern Synthesis. *Biosemiotics*, *14*(1), 5–24. doi: 10.1007/s12304-021-09405-3

Nuño de la Rosa, L. (2010). Becoming organisms: The organisation of development and the development of organisation. *History and Philosophy of the Life Sciences*, *32*(2-3), 289–315. doi: 10.2307/23335076

Nuño de la Rosa, L., & Müller, G. B. (2021). *Evolutionary Developmental Biology: A Reference Guide*. Cham: Springer.

Odling-Smee, J., Laland, K. N., & Feldman, M. W. (2003). *Niche Construction: The Neglected Process in Evolution*. Princeton, NJ: Princeton University Press.

Okasha, S. (2005). On niche construction and extended evolutionary theory. *Biology & Philosophy*, *20*(1), 1–10. doi: 10.1007/s10539-005-0431-3

Okasha, S. (2016). *Philosophy of Science: A Very Short Introduction* (2nd ed.). Oxford: Oxford University Press.

Okasha, S. (2018). *Agents and Goals in Evolution*. Oxford: Oxford University Press.

Ostachuk, A. (2020). The organism and its *Umwelt*: A counterpoint between the theories of Uexküll, Goldstein and Canguilhem. In F. Michelini & K. Köchy (Eds.), *Jakob von Uexküll and Philosophy: Life, Environments, Anthropology* (pp. 158–171). London: Routledge.

Oyama, S. (2000a). Causal democracy and causal contributions in developmental systems theory. *Philosophy of Science*, *67*(3), S332–S347. doi: 10.1086/392830

Oyama, S. (2000b). *The Ontogeny of Information. Developmental Systems and Evolution* (2nd ed.). Durham, NC: Duke University Press.

Oyama, S. (2001). Terms in tension: What do you do when all the good words are taken? In S. Oyama, R. Gray, & P. E. Griffiths (Eds.), *Cycles of Contingency: Developmental Systems and Evolution* (pp. 177–193). Cambridge, MA: The MIT Press.

Oyama, S. (2010). Biologists behaving badly: Vitalism and the language of language. *History and Philosophy of the Life Sciences*, *32*(2-3), 401–423. doi: 10.2307/23335080

Oyama, S., Gray, R., & Griffiths, P. E. (Eds.). (2001). *Cycles of Contingency: Developmental Systems and Evolution*. Cambridge, MA: The MIT Press.

Papineau, D. (1984). Representation and explanation. *Philosophy of Science*, *51*(4), 550–572. doi: 10.1086/289205

Papineau, D. (1987). *Reality and Representation*. Oxford: Basil Blackwell.

Papineau, D. (1993). *Philosophical Naturalism*. Oxford: Basil Blackwell.

Papineau, D. (1998). Teleosemantics and indeterminacy. *Australasian Journal of*

Philosophy, *76*(1), 1–14. doi: 10.1080/00048409812348151

Papineau, D. (2017). Teleosemantics. In D. L. Smith (Ed.), *How Biology Shapes Philosophy. New Foundations for Naturalism* (pp. 95–120). Cambridge: Cambridge University Press.

Pattee, H. H. (2007). The necessity of biosemiotics: Matter-symbol complementarity. In M. Barbieri (Ed.), *Introduction to Biosemiotics: The New Biological Synthesis* (pp. 115–132). Dordrecht: Springer.

Pattee, H. H. (2012). Cell psychology: An evolutionary approach to the symbol-matter problem. In H. H. Pattee & J. Rączaszek-Leonardi (Eds.), *Laws, Language and Life: Howard Pattee's Classic Papers on the Physics of Symbols with Contemporary Commentary* (pp. 165–179). Dordrecht: Springer.

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference* (2nd ed.). New York: Cambridge University Press.

Pearl, J., & Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*. New York: Basic Books.

Pence, C. H. (2021). *The Causal Structure of Natural Selection*. Cambridge: Cambridge University Press.

Pence, C. H., & Ramsey, G. (2013). A new foundation for the propensity interpretation of fitness. *The British Journal for the Philosophy of Science*, *64*(4), 851–881. doi: 10.1093/bjps/axs037

Perbal, L. (2015). The case of the gene: Postgenomics between modernity and postmodernity. *EMBO reports*, *16*(7), 777–781. doi: 10.15252/embr.201540179

Peterson, E. L. (2017). *The Life Organic: The Theoretical Biology Club and the Roots of Epigenetics*. Pittsburgh, PA: University of Pittsburgh Press.

Pharoah, M. (2020). Causation and information: Where is biological meaning to be found? *Biosemiotics*, *13*(3), 309–326. doi: 10.1007/s12304-020-09397-6

Piaget, J. (1971). *Biology and Knowledge: An Essay on the Relations Between Organic Regulations and Cognitive Processes*. Chicago, IL: University of Chicago Press.

Pigliucci, M. (2007). Do we need an extended evolutionary synthesis? *Evolution: International Journal of Organic Evolution*, *61*(12), 2743–2749. doi: 10.1111/j.1558-5646.2007.00246.x

Pigliucci, M. (2010). Genotype–phenotype mapping and the end of the 'genes as blueprint' metaphor. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1540), 557–566. doi: 10.1098/rstb.2009.0241

Pigliucci, M., & Müller, G. B. (Eds.). (2010). *Evolution—The Extended Synthesis*. Cambridge, MA: The MIT Press.

Pinker, S. (1994). *The Language Instinct: How the Mind Creates Language*. New York: William Morrow and Company.

Pinker, S. (1997). *How the Mind Works*. New York: W. W. Norton & Company.

Pinto-Correia, C. (1997). *The Ovary of Eve: Egg and Sperm and Preformation*. Chicago, IL: The University of Chicago Press.

Pittendrigh, C. S. (1958). Adaptation, natural selection, and behavior. In A. Roe & G. G. Simpson (Eds.), *Behavior and Evolution* (pp. 390–416). New Haven, CT: Yale University Press.

Ploeger, A., & Galis, F. (2011). Evo devo and cognitive science. *Wiley Interdisciplinary Reviews. Cognitive Science*, *2*(4), 429–440. doi: 10.1002/wcs.137

Plotkin, H. (1994). *Darwin Machines and the Nature of Knowledge*. London: Allen Lane.

Plotkin, H. (2008). The central problem of cognitive science: The rationalist–empiricist divide. *The Journal of Mind and Behavior*, *29*(1-2), 1–16.

Popper, K. R., & Eccles, J. C. (1983). *The Self and Its Brain: An Argument for Interactionism*. London: Routledge.

Potochnik, A. (2017). *Idealization and the Aims of Science*. Chicago, IL: The University of Chicago Press.

Price, T. D., Qvarnström, A., & Irwin, D. E. (2003). The role of phenotypic plasticity in driving genetic evolution. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1523), 1433–1440. doi: 10.1098/rspb.2003.2372

Prinz, J. J. (2002). *Furnishing the Mind: Concepts and Their Perceptual Basis*. Cambridge, MA: The MIT Press.

Quarfood, M. (2004). *Transcendental idealism and the organism: Essays on kant* (Doctoral dissertation, Department of Philosophy, Stockholm University). Retrieved from `https://www.diva-portal.org/smash/get/diva2:191805/FULLTEXT01.pdf`

Quarfood, M. (2006). Kant on biological teleology: Towards a two-level interpretation. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *37*(4), 735–747. doi: 10.1016/j.shpsc.2006.09.007

Quine, W. V. O. (1976). Two dogmas of empiricism. In *Can Theories Be Refuted?* (pp. 41–64). Springer.

Rama, T. (2018). *Hacia una Psicobiología del Desarrollo para la construcción de Representaciones Conceptuales* (Unpublished master's thesis). Universitat Autònoma de Barcelona, Bellaterra (Barcelona).

Rama, T. (2021). Biosemiotics at the bridge between Eco-Devo and representational theories of mind. *Rivista Italiana di Filosofia del Linguaggio*, *15*(2), 59–92. doi: 10.4396/2021203

Rama, T. (in press). Evolutionary causation and teleosemantics. In J. M. Viejo, M. San Juan, & C. Bueno (Eds.), *Life and Mind — New directions in the*

Philosophy of Biology and Cognitive Sciences. Interdisciplinary Evolution Research - Springer Nature.

Ramsey, G. (2016). The causal structure of evolutionary theory. *Australasian Journal of Philosophy*, *94*(3), 421–434. doi: 10.1080/00048402.2015.1111398

Reid, R. G. B. (2007). *Biological Emergences: Evolution by Natural Experiment.* Cambridge, MA: The MIT Press.

Reisman, K., & Forber, P. (2005). Manipulation and the causes of evolution. *Philosophy of Science*, *72*(5), 1113–1123. doi: 10.1086/508120

Rey, G. (2014). Innate *and* learned: Carey, Mad Dog Nativism, and the poverty of stimuli and analogies (yet again). *Mind & Language*, *29*(2), 109–132. doi: 10.1111/mila.12044

Rheinberger, H.-J., & Müller-Wille, S. (2018). *The Gene: From Genetics to Postgenomics.* Chicago, IL: The University of Chicago Press.

Richards, R. J. (2002). *The Romantic Conception of Life: Science and Philosophy in the Age of Goethe.* Chicago, IL: The University of Chicago Press.

Richardson, S. S., & Stevens, H. (Eds.). (2015). *Postgenomics: Perspectives on Biology after the Genome.* Durham, NC: Duke University Press.

Richerson, P. J., & Boyd, R. (1978). A dual inheritance model of the human evolutionary process i: Basic postulates and a simple model. *Journal of Social and Biological Structures*, *1*(2), 127–154. doi: 10.1016/S0140-1750(78)80002-5

Richerson, P. J., & Boyd, R. (2005). *Not by Genes Alone. How Culture Transformed Human Evolution.* Chicago, IL: The University of Chicago Press.

Ridley, M. (2003). *Evolution* (3rd ed.). Malden, MA: Wiley.

Ritter, W. E. (1919a). *The Unity of the Organism; or, The Organismal Conception of Life* (Vol. 1). Boston, MA: R. G. Badger.

Ritter, W. E. (1919b). *The Unity of the Organism; or, The Organismal Conception of Life* (Vol. 2). Boston, MA: R. G. Badger.

Robert, J. S. (2004). *Embryology, Epigenesis and Evolution: Taking Development Seriously.* Cambridge: Cambridge University Press.

Roli, A., Jaeger, J., & Kauffman, S. A. (2022). How organisms come to know the world: Fundamental limits on artificial general intelligence. *Frontiers in Ecology and Evolution*, 1–14. doi: 10.3389/fevo.2021.806283

Rosen, R. (2000). *Essays on Life Itself.* New York: Columbia University Press.

Rosenberg, A. (1997). Reductionism redux: Computing the embryo. *Biology & Philosophy*, *12*(4), 445–470. doi: 10.1023/A:1006574719901

Rosenberg, A. (2006). *Darwinian Reductionism; or, How to Stop Worrying and Love Molecular Biology.* Chicago, IL: The University of Chicago Press.

Rosenberg, A. (2011). *The Atheist's Guide to Reality: Enjoying Life without Illusions.* New York: W. W. Norton & Company.

Rosenberg, A. (2014). Disenchanted naturalism. In B. Bashour & H. D. Muller (Eds.), *Contemporary Philosophical Naturalism and Its Implications* (pp. 17–36). London: Routledge.

Rosenberg, A. (2020). *Reduction and Mechanism*. Cambridge: Cambridge University Press.

Rosenblueth, A., Wiener, N., & Bigelow, J. (1943). Behavior, purpose and teleology. *Philosophy of Science*, *10*(1), 18–24. doi: 10.1086/286788

Ruiz-Mirazo, K., & Moreno, A. (2004). Basic autonomy as a fundamental step in the synthesis of life. *Artificial Life*, *10*(3), 235–259. doi: 10.1162/1064546041255584

Ruiz-Mirazo, K., Peretó, J., & Moreno, A. (2004). A universal definition of life: Autonomy and open-ended evolution. *Origins of Life and Evolution of the Biosphere*, *34*(3), 323–346. doi: 10.1023/B:ORIG.0000016440.53346.dc

Ruse, M. (1973). Teleological explanations and the animal world. *Mind*, *82*(327), 433–436. doi: 10.1093/mind/LXXXII.327.433

Ruse, M. (2000). Teleology: Yesterday, today, and tomorrow? *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *31*(1), 213–232. doi: 10.1016/S1369-8486(99)00046-1

Ruse, M. (2003). *Darwin and Design: Does Evolution Have a Purpose?* Cambridge, MA: Harvard University Press.

Ruse, M. (2019). *The Darwinian Revolution*. Cambridge: Cambridge University Press.

Russell, B. (1905). On Denoting. *Mind*, *14*(56), 479–493. doi: 10.1093/mind/XIV.4.479

Russell, B. (1910). Knowledge by acquaintance and knowledge by description. *Proceedings of the Aristotelian Society*, *11*, 108–128.

Russell, B. (2010). *The Philosophy of Logical Atomism*. London: Routledge.

Russell, E. S. (1924). *The Study of Living Things: Prolegomena to a Functional Biology*. Methuen & Company Limited.

Russell, E. S. (1930). *The Interpretation of Development and Heredity: A Study in Biological Method*. Oxford: Clarendon Press.

Russell, E. S. (1934). The study of behaviour. *Nature*, *134*(3396), 835–839. doi: 10.1038/134835a0

Russell, E. S. (1945). *The Directiveness of Organic Activities*. Cambridge: Cambridge University Press.

Ryle, G. (1949). *The Concept of Mind*. London: Hutchinson.

Saborido, C., Mossio, M., & Moreno, A. (2011). Biological organization and cross-generation functions. *The British Journal for the Philosophy of Science*, *62*(3), 583–606. doi: 10.1093/bjps/axq034

Saetzler, K., Sonnenschein, C., & Soto, A. M. (2011). Systems biology beyond networks: Generating order from disorder through self-organization. *Seminars in Cancer Biology*, *21*(3), 165–174. doi: 10.1016/j.semcancer.2011.04.004

Salmon, W. C. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press.

Salmon, W. C. (1998). *Causality and Explanation*. Oxford: Oxford University Press.

Salthe, S. N. (1993). *Development and Evolution: Complexity and Change in Biology*. Cambridge, MA: The MIT Press.

Sarkar, S. (1999). From the Reaktionsnorm to the adaptive norm: The norm of reaction, 1909–1960. *Biology & Philosophy*, *14*(2), 235–252. doi: 10.1023/A:1006690502648

Sarkar, S. (2005). *Molecular Models of Life: Philosophical Papers on Molecular Biology*. Cambridge, MA: The MIT Press.

Sarkar, S. (2006). From genes as determinants to DNA as resource: Historical notes on development and genetics. In E. M. Neumann-Held & C. Rehmann-Sutter (Eds.), *Genes in Development: Re-reading the Molecular Paradigm* (pp. 77–95). Durham, NC: Duke University Press.

Scarantino, A. (2015). Information as a probabilistic difference maker. *Australasian Journal of Philosophy*, *93*(3), 419–443. doi: 10.1080/00048402.2014.993665

Schlichting, C. D., & Pigliucci, M. (1998). *Phenotypic Evolution: A Reaction Norm Perspective*. Sunderland, MA: Sinauer.

Schlosser, G. (1998). Self-re-production and functionality. *Synthese*, *116*(3), 303–354. doi: 10.1023/A:1005073307193

Schmalhausen, I. I. (1949). *Factors of Evolution: The Theory of Stabilizing Selection*. Philadelphia, PA: Blakiston.

Schneirla, T. C. (1966). Behavioral development and comparative psychology. *The Quarterly Review of Biology*, *41*(3), 283–302. doi: 10.1086/405056

Schrödinger, E. (1944). *What is life? The Physical Aspect of the Living Cell*. Cambridge: Cambridge University Press.

Schulte, P. (2015). Perceptual representations: A teleosemantic answer to the breadth-of-application problem. *Biology & Philosophy*, *30*(1), 119–136. doi: 10.1007/s10539-013-9390-2

Schulte, P. (2018). Perceiving the world outside: How to solve the distality problem for informational teleosemantics. *The Philosophical Quarterly*, *68*(271), 349–369. doi: 10.1093/pq/pqx052

Schulte, P. (2020). Why mental content is not like water: Reconsidering the reductive claims of teleosemantics. *Synthese*, *197*(5), 2271–2290. doi: 10.1007/s11229-018-1808-6

Scott-Phillips, T. C., Laland, K. N., Shuker, D. M., Dickins, T. E., & West, S. A.

(2014). The niche construction perspective: A critical appraisal. *Evolution*, *68*(5), 1231–1243. doi: 10.1111/evo.12332

Searle, J. R. (1992). *The Rediscovery of the Mind*. Cambridge, MA: The MIT press.

Searle, J. R. (1995). *The Construction of Social Reality*. New York: Free Press.

Sebeok, T. A. (1988). 'Animal' in biological and semiotic perspective. In T. Ingold (Ed.), *What is an Animal?* (pp. 63–76). London: Routledge.

Sebeok, T. A. (2001). Biosemiotics: Its roots, proliferation, and prospects. *Semiotica*, *134*(1-4), 61–78. doi: 10.1515/semi.2001.014

Shannon, C. E., & Weaver, W. (1949). *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.

Shea, N. (2007a). Consumers need information: Supplementing teleosemantics with an input condition. *Philosophy and Phenomenological Research*, *75*(2), 404–435. doi: https://doi.org/10.1111/j.1933-1592.2007.00082.x

Shea, N. (2007b). Representation in the genome and in other inheritance systems. *Biology & Philosophy*, *22*(3), 313–331. doi: https://doi.org/10.1007/s10539-006-9046-6

Shea, N. (2011). Developmental Systems Theory formulated as a claim about inherited representations. *Philosophy of Science*, *78*(1), 60–82. doi: https://doi.org/10.1086/658110

Shea, N. (2012a). Genetic representation explains the cluster of innateness-related properties. *Mind & Language*, *27*(4), 466–493. doi: https://doi.org/10.1111/j.1468-0017.2012.01452.x

Shea, N. (2012b). New thinking, innateness and inherited representation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1599), 2234–2244. doi: 0.1098/rstb.2012.0125

Shea, N. (2013). Inherited representations are read in development. *The British Journal for the Philosophy of Science*, *64*(1), 1–31. doi: https://doi.org/10.1093/bjps/axr050

Shea, N. (2018). *Representation in Cognitive Science*. Oxford: Oxford University Press.

Simpson, G. G. (1944). *Tempo and Mode in Evolution*. New York: Columbia University Press.

Simpson, G. G. (1953). The Baldwin Effect. *Evolution. International Journal of Organic Evolution*, *7*(2), 110–117. doi: 10.1111/j.1558-5646.1953.tb00069.x

Skinner, B. F. (1953). *Science and Human Behavior*. New York: Macmillan.

Skinner, B. F. (1957). *Verbal Bahavior*. New York: Appleton-Century-Crofts.

Skyrms, B. (2010). *Signals: Evolution, Learning, and Information*. Oxford: Oxford University Press.

Smolin, L. (2001). *Three Roads to Quantum Gravity*. New York: Basic Books.

Soames, S. (2010). *Philosophy of Language*. Princeton, NJ: Princeton University Press.

Sober, E. (1980). Evolution, population thinking, and essentialism. *Philosophy of Science*, *47*(3), 350–383. doi: 10.1086/288942

Sober, E. (1984). *The Nature of Selection. Evolutionary Theory in Philosophical Focus*. Chicago, IL: The University of Chicago Press.

Sober, E. (Ed.). (2006). *Conceptual Issues in Evolutionary Biology* (3rd ed.). Cambridge, MA: The MIT Press.

Sober, E. (2013). Trait fitness is not a propensity, but fitness variation is. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *44*(3), 336–341. doi: 10.1016/j.shpsc.2013.03.002

Spelke, E. S. (1998). Nativism, empiricism, and the origins of knowledge. *Infant Behavior and Development*, *21*(2), 181–200. doi: 10.1016/s0163-6383(98)90002-9

Spelke, E. S. (2017). Core knowledge, language, and number. *Language Learning and Development*, *13*(2), 147–170. doi: 10.1080/15475441.2016.1263572

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*(1), 89–96. doi: 10.1111/j.1467-7687.2007.00569.x

Sperber, D. (1996). *Explaining Culture. A Naturalistic Approach*. Oxford: Blackwell.

Stampe, D. W. (1977). Toward a causal theory of linguistic representation. *Midwest Studies in Philosophy*, *2*(1), 42–63. doi: 10.1111/j.1475-4975.1977.tb00027.x

Stegmann, U. (2017). Evolution and information: an overview. In *The Routledge Handbook of Evolution and Philosophy* (pp. 79–90). New York: Routledge. doi: 10.4324/9781315764863.ch6

Stephens, C. (2004). Selection, drift, and the "forces" of evolution. *Philosophy of Science*, *71*(4), 550–570. doi: 10.1086/423751

Sterelny, K. (2010). Minds: Extended or scaffolded? *Phenomenology and the Cognitive Sciences*, *9*(4), 465–481. doi: 10.1007/s11097-010-9174-y

Sterelny, K., & Griffiths, P. E. (1999). *Sex and Death: An Introduction to Philosophy of Biology*. Chicago, IL: The University of Chicago Press.

Sterelny, K., Smith, K. C., & Dickison, M. (1996). The extended replicator. *Biology & Philosophy*, *11*(3), 377–403. doi: 10.1007/BF00128788

Stotz, K. (2006a). Molecular epigenesis: Distributed specificity as a break in the Central Dogma. *History and Philosophy of the Life Sciences*, *28*(4), 533–548.

Stotz, K. (2006b). With 'genes' like that, who needs an environment? Postgenomics's argument for the 'Ontogeny of Information'. *Philosophy of Science*, *73*(5), 905–917. doi: 10.1086/518748

Stotz, K. (2008). The ingredients for a postgenomic synthesis of nature and nurture. *Philosophical Psychology*, *21*(3), 359–381. doi: 10.1080/09515080802200981

Stotz, K. (2010). Human nature and cognitive–developmental niche construction. *Phenomenology and the Cognitive Sciences*, *9*(4), 483–501. doi: 10.1007/s11097-010-9178-7

Stotz, K. (2014). Extended evolutionary psychology: The importance of transgenerational developmental plasticity. *Frontiers in Psychology*, *5*, 908. doi: 10.3389/fpsyg.2014.00908

Stotz, K. (2017). Why developmental niche construction is not selective niche construction: And why it matters. *Interface Focus*, *7*(5), 20160157. doi: 10.1098/rsfs.2016.0157

Stotz, K. (2019). Biological information in developmental and evolutionary systems. In T. Uller & K. L. Laland (Eds.), *Evolutionary Causation: Biological and Philosophical Reflections* (pp. 323–344). Cambridge, MA: The MIT Press.

Stotz, K., Adam Bostanci, A., & Griffiths, P. E. (2006). Tracking the shift to 'Postgenomics'. *Public Health Genomics*, *9*(3), 190–196. doi: 10.1159/000092656

Stotz, K., & Griffiths, P. (2016). Epigenetics: Ambiguities and implications. *History and Philosophy of the Life Sciences*, *38*(4), 1–20. doi: 10.1007/s40656-016-0121-2

Stotz, K., & Griffiths, P. E. (2017). Biological information, causality and specificity: An intimate relationship. In S. I. Walker, P. C. W. Davies, & G. F. R. Ellis (Eds.), *From Matter to Life. Information and Causality* (pp. 366–390). Cambridge: Cambridge University Press.

Sultan, S. E. (2004). Promising directions in plant phenotypic plasticity. *Perspectives in Plant Ecology, Evolution and Systematics*, *6*(4), 227–233. doi: 10.1078/1433-8319-00082

Sultan, S. E. (2010). Plant developmental responses to the environment: Eco-devo insights. *Current Opinion in Plant Biology*, *13*(1), 96–101. doi: 10.1016/j.pbi.2009.09.021

Sultan, S. E. (2015). *Organism & Environment: Ecological Development, Niche Construction, and Adaption*. Oxford: Oxford University Press.

Sultan, S. E. (2019). Genotype-environment interaction and the unscripted reaction norm. In T. Uller & K. N. Laland (Eds.), *Evolutionary Causation: Biological and Philosophical Reflections* (pp. 109–126). Cambridge, MA: The MIT Press.

Sultan, S. E. (2021). Phenotypic plasticity as an intrinsic property of organisms. In D. W. Pfennig (Ed.), *Phenotypic Plasticity & Evolution: Causes,*

Consequences, Controversies (pp. 3–24). Boca Raton, FL: CRC Press.

Sultan, S. E., Moczek, A. P., & Walsh, D. (2022). Bridging the explanatory gaps: What can we learn from a biological agency perspective? *BioEssays*, *44*(1), 2100185. doi: 10.1002/bies.202100185

Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: Harvard University Press.

Thorner, J., Hunter, T., Cantley, L. C., & Sever, R. (2014). Signal transduction: From the atomic age to the post-genomic era. *Cold Spring Harbor Perspectives in Biology*, *6*(12), a022913. doi: 10.1101/cshperspect.a022913

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambrdige; MA: Harvard University Press.

von Uexküll, J. (1923). Weltanschauung und Gewissen. *Deutsche Rundschau*, *197*, 253–266.

von Uexküll, J. (1926). *Theoretical Biology*. London: Kegan Paul, Trench, Trubner & Co.

von Uexküll, J. (1936). *Niegeschaute Welten. Die Umwelte meiner Freunde*. Berlin: S. Fischer Verlag. (Excerpt selected and translated in von Uexküll, 2001)

von Uexküll, J. (2001). An introduction to Umwelt. *Semiotica*, *134*(1-4). doi: 10.1515/semi.2001.017

von Uexküll, J. (2010). *A Foray Into the Worlds of Animals and Humans; with A Theory of Meaning*. Minneapolis, MN: University of Minnesota Press.

Uller, T. (2008). Developmental plasticity and the evolution of parental effects. *Trends in Ecology & Evolution*, *23*(8), 432–438. doi: 10.1016/j.tree.2008.04.005

Uller, T., Feiner, N., Radersma, R., Jackson, I. S. C., & Rago, A. (2020). Developmental plasticity and evolutionary explanations. *Evolution & Development*, *22*(1-2), 47–55. doi: 10.1111/ede.12314

Van den Berg, H. (2013). The Wolffian roots of Kant's teleology. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *44*(4, Part B), 724–734. doi: 10.1016/j.shpsc.2013.07.003

Van Duijn, M., Keijzer, F., & Franken, D. (2006). Principles of minimal cognition: Casting cognition as sensorimotor coordination. *Adaptive Behavior*, *14*(2), 157–170. doi: 10.1177/105971230601400207

Varela, F. J. (1979). *Principles of Biological Autonomy*. New York: North Holland.

Varela, F. J., & Maturana, H. (1972). Mechanism and biological explanation. *Philosophy of Science*, *39*(3), 378–382. doi: 10.1086/288458

Veit, W. (2021). Agential thinking. *Synthese*, *199*(5), 13393–13419. doi: 10.1007/s11229-021-03380-5

Vygotsky, L. S. (1978). *Mind in Society: Development of Higher Psychological*

Processes. Cambridge, MA: Harvard University Press.

Waddington, C. H. (1941). Evolution of developmental systems. *Nature*, *147*, 108–110. doi: 10.1038/147108a0

Waddington, C. H. (1953). Genetic assimilation of an acquired character. *Evolution. International Journal of Organic Evolution*, *7*(2), 118–126. doi: 10.1111/j.1558-5646.1953.tb00070.x

Waddington, C. H. (1957). *The Strategy of Genes*. London: Allen & Unwin.

Waddington, C. H. (1959). Evolutionary systems—Animal and human. *Nature*, *183*(4676), 1634–1638. doi: 10.1038/1831634a0

Wagner, A. (2013). *Robustness and Evolvability in Living systems*. Princeton, NJ: Princeton University Press.

Wagner, G. P. (Ed.). (2001). *The Character Concept in Evolutionary Biology*. San Diego, CA: Academic Press.

Wagner, G. P. (2014). *Homology, Genes, and Evolutionary Innovation*. Princeton, NJ: Princeton University Press.

Wagner, G. P. (2016). What is "homology thinking" and what is it for? *Journal of Experimental Zoology Part B (Molecular and Developmental Evolution)*, *326*(1), 3–8. doi: 10.1002/jez.b.22656

Wagner, G. P., & Larsson, H. C. E. (2003). What is the promise of developmental evolution? III. The crucible of developmental evolution. *Journal of Experimental Zoology Part B (Molecular and Developmental Evolution)*, *300B*(1), 1–4. doi: 10.1002/jez.b.41

Walsh, D. M. (2000). Chasing shadows: Natural selection and adaptation. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *31*(1), 135–153. doi: 10.1016/S1369-8486(99)00041-2

Walsh, D. M. (2003). Fit and diversity: Explaining adaptive evolution. *Philosophy of Science*, *70*(2), 280–301. doi: 10.1086/375468

Walsh, D. M. (2006a). Evolutionary essentialism. *The British Journal for the Philosophy of Science*, *57*(2), 425–448. doi: 10.1093/bjps/axl001

Walsh, D. M. (2006b). Organisms as natural purposes: The contemporary evolutionary perspective. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *37*(4), 771–791. doi: 10.1016/j.shpsc.2006.09.009

Walsh, D. M. (2007a). Development: Three grades of ontogenetic involvement. In M. Matthen & C. Stephens (Eds.), *Handbook of the Philosophy of Science* (Vol. 3: Philosophy of Biology, pp. 179–199). Elsevier.

Walsh, D. M. (2007b). The pomp of superfluous causes: The interpretation of evolutionary theory. *Philosophy of Science*, *74*(3), 281–303. doi: 10.1086/520777

Walsh, D. M. (2008). Teleology. In M. Ruse (Ed.), *The Oxford Handbook of Philosophy of Biology* (pp. 113–137). Oxford: Oxford University Press.

Walsh, D. M. (2010). Two neo-Darwinisms. *History and Philosophy of the Life Sciences*, *32*(2-3), 317–339.

Walsh, D. M. (2012a). Mechanism and purpose: A case for natural teleology. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *43*(1), 173–181. doi: 10.1016/j.shpsc.2011.05.016

Walsh, D. M. (2012b). Situated aaptationism. In W. P. Kabasenche, M. O'Rourke, & M. H. Slater (Eds.), *The Environment: Philosophy, Science, and Ethics* (pp. 89–116). Cambridge, MA: The MIT Press.

Walsh, D. M. (2013a). The affordance landscape: The spatial metaphors of evolution. In G. Barker, E. Desjardins, & T. Pearce (Eds.), *Entangled Life. Organism and Environment in the Biological and Social Sciences* (pp. 213–236). Dordrecht: Springer.

Walsh, D. M. (2013b). Mechanism, emergence, and miscibility: The autonomy of evo-devo. In P. Huneman (Ed.), *Functions: Selection and Mechanisms* (pp. 43–65). Dordrecht: Springer.

Walsh, D. M. (2013c). The negotiated organism: Inheritance, development, and the method of difference. *Biological Journal of the Linnean Society*, *112*(2), 295–305. doi: 10.1111/bij.12118

Walsh, D. M. (2014). Function and teleology. In R. P. Thompson & D. M. Walsh (Eds.), *Evolutionary Biology. Conceptual, Ethical, and Religious Issues* (pp. 193–216). Cambridge: Cambridge University Press.

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

Walsh, D. M. (2018). Objectcy and agency: Towards a methodological vitalism. In D. J. Nicholson & J. Dupré (Eds.), *Everything Flows. Towards a Processual Philosophy of Biology* (pp. 167–185). Oxford: Oxford University Press.

Walsh, D. M. (2019). The paradox of population thinking: First order causes and higher order effects. In T. Uller & K. N. Laland (Eds.), *Evolutionary Causation: Biological and Philosophical Reflections* (pp. 227–246). Cambridge, MA: The MIT Press.

Walsh, D. M. (2021). Teleology in evo-devo. In L. Nuño de la Rosa & G. B. Müller (Eds.), *Evolutionary Developmental Biology: A Reference Guide* (pp. 495–508). Springer.

Walsh, D. M., Ariew, A., & Matthen, M. (2017). Four pillars of statisticalism. *Philosophy, Theory, and Practice in Biology*, *9*(20171201). doi: 10.3998/ptb.6959004.0009.001

Walsh, D. M., Lewens, T., & Ariew, A. (2002). The trials of life: Natural selection

and random drift. *Philosophy of Science*, *69*(3), 429–446. doi: 10.1086/342454

Watson, J. D., & Crick, F. H. C. (1953). Molecular structure of nucleic acids: A structure for Deoxyribose Nucleic Acid. *Nature*, *171*(4356), 737–738. doi: 10.1038/171737a0

Weber, A., & Varela, F. J. (2002). Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, *1*(2), 97–125. doi: 10.1023/a:1020368120174

Weber, M. (2022). *Philosophy of Developmental Biology*. Cambridge: Cambridge University Press.

Weismann, A. (1892). *Das Keimplasma. Eine Theorie der Vererbung*. Jena: Gustav Fischer.

Weiss, P. A. (1939). *Principles of Development*. New York: Holt, Rinehart, and Winston.

West, M. J., & King, A. P. (1987). Settling nature and nurture into an ontogenetic niche. *Developmental Psychobiology*, *20*(5), 549–562. doi: 10.1002/dev.420200508

West, M. J., King, A. P., & White, D. J. (2003). The case for developmental ecology. *Animal Behaviour*, *66*(4), 617–622. doi: 10.1006/anbe.2003.2221

West-Eberhard, M. J. (2003). *Developmental Plasticity and Evolution*. Oxford: Oxford University Press.

West-Eberhard, M. J. (2005). Developmental plasticity and the origin of species differences. *Proceedings of the National Academy of Sciences*, *102*(suppl 1), 6543–6549. doi: 10.1073/pnas.050184410

Wheeler, M. (2011). Mind in life or life in mind? Making sense of deep continuity. *Journal of Consciousness Studies*, *18*(5-6), 148–168.

Why biosemiotics? An introduction to our view on the biology of life itself. (n.d.).

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

Wiese, W., & Friston, K. J. (2021). Examining the continuity between life and mind: Is there a continuity between autopoietic intentionality and representationality? *Philosophies*, *6*(1), 18. doi: 10.3390/philosophies6010018

Williams, G. C. (1966). *Adaptation and Natural Selection. A Critique of Some Current Evolutionary Thought*. Princeton, NJ: Princeton University Press.

Williams, P. L., & Beer, R. D. (2010). *Nonnegative decomposition of multivariate information*. Retrieved from `arXiv:1004.2515` doi: 10.48550/arXiv.1004.2515

Wilson, E. O. (1975). *Sociobiology: The New Synthesis*. Cambridge, MA: Harvard University Press.

Wimsatt, W. C. (1986). Developmental constraints, generative entrenchment, and

the innate-acquired distinction. In W. Bechtel (Ed.), *Integrating Scientific Disciplines* (pp. 185–208). Dordrecht: Martinus Nijhoff.

Wittgenstein, L. (2010). *Philosophical investigations*. Mahwa, NJ: Wiley-Blackwell.

Woese, C. R. (2004). A new biology for a new century. *Microbiology and Molecular Biology Reviews*, *68*(2), 173–186. doi: 10.1128/MMBR.68.2.173-186.2004

Wolpert, L. (1994). Do we understand development? *Science*, *266*(5185), 571–573. doi: 10.1126/science.7939707

Woltereck, R. (1909). Weitere experimentelle Untersuchungen über Artveränderung, speziell über das Wesen quantitativer Artunterschiede bei Daphniden. *Verhandlungen der Deutschen Zoologischen Gesellschaft*, *19*, 110–172.

Woodger, J. H. (1929). *Biological Principles: A Critical Study*. London: Kegan Paul and Co.

Woodward, J. (2002). What is a mechanism? A counterfactual account. *Philosophy of Science*, *69*(S3), S366–S377. doi: 10.1086/341859

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.

Woodward, J. (2010). Causation in biology: Stability, specificity, and the choice of levels of explanation. *Biology & Philosophy*, *25*(3), 287–318. doi: 10.1007/s10539-010-9200-z

Wright, L. (1976). *Teleological Explanations: An Etiological Analysis of Goals and Functions*. Berkeley, CA: University of California Press.

Wund, M. A. (2012). Assessing the impacts of phenotypic plasticity on evolution. *Integrative and Comparative Biology*, *52*(1), 5–15. doi: 10.1093/icb/ics050

Zammito, J. H. (2006). Teleology then and now: The question of Kant's relevance for contemporary controversies over function in biology. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *37*(4), 748–770. doi: 10.1016/j.shpsc.2006.09.008

Zammito, J. H. (2012). The Lenoir thesis revisited: Blumenbach and Kant. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *43*(1), 120–132. doi: 10.1016/j.shpsc.2011.05.011