UNIVERSIDAD DE MURCIA

ESCUELA INTERNACIONAL DE DOCTORADO

Questions at the Intersection between Metaphysics and Biology: Towards a Metaphysics of Biology

Cuestiones en la Intersección entre Metafísica y Biología: Hacia una Metafísica de la Biología

Dña. Vanesa Triviño Alonso
2019

UNIVERSIDAD DE MURCIA

FACULTAD DE FILOSOFÍA

Questions at the Intersection between Metaphysics and Biology: Towards a Metaphysics of Biology

Cuestiones en la Intersección entre Metafísica y Biología: Hacia una Metafísica de la Biología

Vanessa Triviño Alonso

DIRECTORAS
María Cerezo Lallana (Universidad de Murcia)
Laura Nuño de la Rosa García (Universidad Complutense de Madrid)

2019

A Carlos, por ser lo que siempre permanece
mientras todo lo demás cambia.

Acknowledgments

Due to all the people I have met during the whole research process starting in 2013, as well as the constant conversations and debates I have had with my supervisors, colleagues, and professors of my Faculty, I do not consider this thesis to be fully mine. It is due to all the support and feedback I have received that I have been able to think of these questions in the way I do. I thank the Ph.D. students of the faculty of Philosophy that have shared with me the personal and intellectual process of writing a thesis, and with whom I had the opportunity to share frustrations, tips, advice, and feedback: Jesús López Campillo, Agustina Varela, Rafael Pérez Barquero, Nkiru Onwuasoanya, and Pablo Domenech. I also thank the professors of my faculty for listening to some of my initial ideas in different seminars, and for all the resources they have taught me in order to approach philosophical questions: Jorge Novella, Gustavo Fernández, Ángel García, María José Alcaraz, and Francisca Pérez. I am particularly grateful to Paco Calvo, who gave me feedback to the initial draft of my paper on fitness, making it improve qualitatively. Furthermore, the passion he transmits in his classes when presenting philosophical ideas, as well as the clear way in which he does it, has always encouraged me to improve my philosophical reasoning and to work harder and better, especially in thinking of possible anti-reductionist positions that might convince him.

My greatest debt is to my supervisors, who share with me the interest in both metaphysics and philosophy of biology and that have been by my side during the whole process of conception and development of this dissertation, encouraging me to always improve and be a better researcher. I am very fortunate in working with them. Without all of their patient and support, I wouldn’t have grown both as a philosopher and as a person in the way I have

done it. Thank you for teaching me how to enjoy doing philosophical research, and the importance of being happy with anything I do both in the personal and in the professional senses. María Cerezo, in teaching and researching with complete passion, has been my main motivation for doing research since my early days as a student of philosophy. Thanks for all your patience when teaching and correcting me, for all the hours we have spent in deliberating about a particular topic, and for always making me realize of the importance of making precise claims and arguments by continuously showing me the undesired implications of my ideas. Laura Nuño de la Rosa has taught me to appreciate the importance of calm reflection concerning philosophical questions in order to develop clear ideas. Thanks for always understanding me and ordering my philosophical ideas when they are a mess in my mind, for showing me the complexity and richness of biology and the important role the philosopher plays in this field, and thank you for all the laughs we have had in exploring possible ideas and doing research.

I would also like to thank my friends and family for the unconditional patience they have had throughout the long and difficult process that has been the conception of this thesis. A mis amigos, Mª Carmen Salazar, Ana Escuer, Sol de la Torre, Rosa Alcaraz, Jorge Hernández, Juan Campos, Maddi Etxabe, Arancha Hervás, y Cristina Soguero. Especial mención merece Cristina Villegas, con quien he tenido la suerte de compartir muchas de las horas de trabajo que supone el realizar una tesis. Gracias por estar conmigo prácticamente desde el principio, has sido mi apoyo incondicional tanto en lo personal como en lo profesional. Gracias también por enriquecerme compartiendo tus pensamientos filosóficos conmigo, y por hacerme ver y cuestionar todas las ideas que muchas veces asumo sin más. Has amenizado este camino tan difícil llenándolo de momentos inolvidables.

A mis padres, por enseñarme el valor del trabajo y el esfuerzo. A mi padre, Francisco, por transmitirme el valor de la curiosidad y la duda, y enseñarme a reflexionar desde que era pequeña a través de cada charla que hemos mantenido. A mi madre, Josefa, por no dejar que me olvidara nunca de las cosas realmente importantes de la vida, y enseñarme a apreciar el valor de todo lo que me rodea. A mi hermana, Raquel, por toda la paciencia ha tenido conmigo, y por entender todo lo que la filosofía y la investigación significan para mí. Gracias por haber sido mi punto de unión con la realidad cada vez que me absorbía el mundo inteligible. A mi primo, Fernando, que con sus llamadas diarias hizo que me sintiera menos sola durante mis estancias. A mi suegro, Antonio, por vivir cada etapa de mi desarrollo académico de cerca, y demostrarme siempre que esta tesis es tan importante para ti como lo puede ser para mí. A mi abuelo, Ángel, porque se fue en el estadio inicial de mi aventura como doctoranda soñando con verme acabar. Estoy segura de que me sigues viendo desde donde estés, por eso hay un pedacito de ti en esta tesis. Y, finalmente, a Carlos, porque ya sabía que sería filósofa mucho antes de que yo me diera cuenta. Quiero agradecerle su apoyo incondicional, su paciencia infinita en mis momentos más inaguantables, y el hecho de anteponer siempre mi sueño de ser doctora a cualquier otra cosa. Tú has sido el verdadero motor que me ha impulsado a seguir adelante cada vez que he querido abandonar. Detrás de cada pequeño paso que he dado y de cada logro que he conseguido, estás tú.

Resumen

La filosofía de la biología comenzó a configurarse como disciplina a finales de la década de los setenta. Debido a la influencia del positivismo lógico, por un lado, y de la síntesis moderna de la evolución, por otro, los filósofos de la biología abogaron por una aproximación metodológica y epistemológica de los problemas conceptuales de la biología, rechazando los enfoques de carácter metafísico. Desde finales de la década de los noventa, sin embargo, esta situación parece estar cambiando, y ya no es extraño encontrar en la actualidad a filósofos de la biología que tienen en consideración teorías y conceptos metafísicos para abordar cuestiones conceptuales de la biología. Trabajos recientes, como el de *Everything Flows: Towards a Processual Philosophy of Biology*, publicado el pasado año por Daniel Nicholson y John Dupré, ilustran esta tendencia. De hecho, algunos filósofos ya han comenzado a hablar de una 'metafísica de la biología' (Guay y Pradeu, 2017).

Esta tesis doctoral aborda la interacción entre metafísica y biología tal y como se despliega en filosofía de la biología, y el uso que se hace de la metafísica en este contexto para abordar los problemas conceptuales de la biología. Con este objetivo, he llevado a cabo un trabajo sistemático de categorización de los trabajos en filosofía de la biología donde aparece esta interacción, y he ejercitado yo misma esta aproximación metafísica a dos problemas conceptuales en biología, a saber: la definición del concepto de aptitud *(fitness)* y el debate sobre el estatus ontológico de los holobiontes. Como se argumenta a lo largo de la tesis, estos ejemplos ilustran de forma clara la interacción entre metafísica y biología.

Mi elección de estos problemas se fundamenta en varias razones. En primer lugar, ambos debates permiten ilustrar la utilidad de la metafísica tanto en cuestiones de filosofía de la biología más clásicas o tradicionales, como el problema de la aptitud, como en cuestiones más contemporáneas tales como el debate sobre el estatus ontológico de los holobiontes. En concreto, en el caso de *fitness,* tanto filósofos de la biología como los propios biólogos han reivindicado la necesidad de profundizar en la caracterización ontológica de esta propiedad. En cuanto al problema de los holobiontes, resulta un problema particularmente atractivo desde una perspectiva metafísica, dado que nociones metafísicas como 'individuo', 'persistencia', 'composición' o '*parteidad*' aparecen recurrentemente en el mismo.

La tesis está articulada en cuatro capítulos. El Capítulo 1 ("Hacia una metafísica de la biología") introduce brevemente el debate en metafísica de la ciencia acerca de la interacción entre metafísica y ciencia para examinar, a continuación, la relación particular entre metafísica y biología. Con este objetivo, se introducen los orígenes de la filosofía de la biología como disciplina y se identifica la tendencia actual de recurrir a la metafísica para abordar filosóficamente problemas conceptuales en biología. El creciente recurso a la metafísica por parte de los filósofos de la biología, así como el debate en metafísica de la ciencia constituyen el marco conceptual de la tesis, a partir del cual se aborda la interacción entre metafísica y biología. En concreto, se proponen dos formas de interacción, a saber: 'metafísica *para* la biología' y 'metafísica *en* la biología'.

El resto de capítulos muestra los resultados obtenidos al abordar los problemas de la aptitud y los holobiontes desde un marco metafísico. En el Capítulo 2, titulado "La aptitud biológica como disposición causal", se introduce la teoría disposicional de la causación, que se emplea como marco conceptual para desarrollar una caracterización de la aptitud como propiedad disposicional de los organismos. En tanto que disposición causal, defiendo que la aptitud resulta de la combinación no lineal de las funciones disposicionales que caracterizan al organismo y que están orientadas hacia un determinado efecto (la supervivencia y la reproducción), el cual se alcanza una vez que la combinación de esas disposiciones excede un determinado umbral. Como se argumenta a lo largo del capítulo, esta caracterización de la aptitud ofrece una serie de ventajas sobre otras caracterizaciones metafísicas ofrecidas con anterioridad, como la que considera la aptitud como propiedad superveniente (Rosenberg, 1978) o la interpretación propensionista (Mills y Beatty, 1979). Por un lado, la caracterización de la aptitud como disposición causal de los organismos, gracias a recursos conceptuales como la noción de 'medio normal' o la distinción entre 'factores que intervienen' y 'factores que previenen', permite distinguir de forma más adecuada entre 'aptitud esperada' y 'aptitud realizada'. Por otro lado, esta caracterización de la aptitud permite dar cuenta del carácter cambiante de esta propiedad de los organismos a lo largo de su vida, así como de su carácter contextual, ya que la aptitud depende del entorno ambiental y poblacional en el que se encuentra el organismo.

El Capítulo 3, titulado "El carácter emergente de la aptitud biológica", introduce la noción metafísica de emergencia, así como las diferentes formas de caracterizar las nociones de dependencia y autonomía asociadas a este concepto. A partir de este marco metafísico, se examina la forma de dependencia y autonomía que se dan en el caso de la aptitud biológica, y se argumenta que esta propiedad es emergente en sentido fuerte. El argumento principal es que el poder causal de la aptitud, a saber, posibilitar que el organismo sea capaz de sobrevivir y reproducirse en un determinado medio y población, no puede reducirse al poder causal de sus funciones disposicionales. En este sentido, argumentamos que la relación entre la aptitud y su base es causal, por un lado, y, por otro, que la aptitud tiene un poder causal autónomo que se manifiesta de forma descendente.

Finalmente, en el Capítulo 4, titulado "Relaciones de *parteidad* en holobiontes", prestamos atención al debate sobre el estatus ontológico de las entidades que resultan de las relaciones simbióticas. Los llamados 'holobiontes' se han interpretado ontológicamente como individuos y como comunidades ecológicas. Recientemente, la llamada teoría eco-inmune de los holobiontes (Chiu y Eberl, 2016) propone una solución metafísica intermedia entre ambos extremos en la que los holobiontes se caracterizan como 'híbridos' entre el hospedador y sus huéspedes microbianos. Este capítulo explora las implicaciones metafísicas que se siguen acerca del concepto de *parteidad* en esta propuesta, argumentando que la caracterización de los holobiontes como híbridos no está metafísicamente fundada. En primer lugar, muestro que parece haber dos nociones diferentes de *parteidad* implícitas en esta teoría que juegan roles diferentes de modo no justificado: la primera noción se aplica a la hora de explicar la relación entre el hospedador y la microbiota; la otra se emplea para explicar la relación del hospedador y sus componentes no microbianos, como sus órganos. En segundo lugar, argumento que extender la noción de *parteidad* empleada en el caso de la microbiota al resto de los componentes del hospedador da lugar a conclusiones contraintuitivas, como la de que los órganos no son partes propias del hospedador.

La aproximación metafísica al problema de la aptitud y del estatus ontológico de los holobiontes contribuyen a la caracterización de la interacción entre metafísica y biología presentada en el Capítulo 1. De hecho, cada problema instancia una interacción distinta entre metafísica y biología. El problema de la aptitud instancia la interacción entre metafísica y biología que se ha denominado "metafísica *para* la biología". Esta interacción tiene lugar cuando la filosofía de la biología apela a teorías y conceptos metafísicos para clarificar el tipo de realidad a la que refieren los conceptos biológicos, contribuyendo a clarificarlos y a resolver problemas asociados a los mismos. Los Capítulos 2 y 3 han procedido de este modo, apelando directamente a teorías y nociones metafísicas como la teoría disposicional de la causación y la noción de emergencia, para clarificar el estatus ontológico de la aptitud biológica. A su vez, la clarificación metafísica de esta noción, entendida como una disposición causal de los organismos que es emergente en sentido fuerte, permite solucionar algunos de los problemas clásicamente asociados a la aptitud, tales como el problema de la circularidad o la falta de criterio para diferenciar entre selección natural y azar.

El problema de los holobiontes, por su parte, instancia la interacción entre metafísica y biología que hemos llamado "metafísica *en* la biología". Esta interacción tiene lugar cuando en filosofía de la biología se atiende a los presupuestos e implicaciones metafísicas implícitos en las distintas teorías, fenómenos y prácticas biológicas. En el Capítulo 4 hemos procedido de este modo, empleando la metafísica de una forma indirecta, con el fin de hacer explícitos presupuestos e implicaciones metafísicos. En este caso, hemos explorado las implicaciones metafísicas implícitas en la noción de *parteidad* que emplea la teoría eco-inmune de los holobiontes. A este respecto, desarrollamos tres argumentos principales. En primer lugar, argumentamos que la integración funcional puede considerarse un criterio de *parteidad*. En segundo lugar, defendemos que las partes propias de un objeto son contingentes, cambiantes y pueden compartirse con otros objetos. Finalmente, argumentamos que la *parteidad* no es un caso especial de andamiaje (*scaffolding*). Al hacer explícitos estos presupuestos metafísicos, esperamos haber contribuido al debate sobre el estatus ontológico de los holobiontes o, al menos, haber mostrado que las razones ofrecidas para defender el carácter híbrido de los holobiontes no están metafísicamente bien fundamentadas, por lo que el debate acerca de si los holobiontes son individuos o comunidades ecológicas sigue aún abierto.

Por otro lado, esta exploración y ejercicio de la interacción entre metafísica y biología también ha prestado atención al tipo de metafísica que desempeña el rol en esta interacción. Como se argumenta en el primer capítulo, esta metafísica no se corresponde con ninguno de los tipos de metafísica que suelen reconocerse en metafísica de la ciencia, a saber: la metafísica a priori y la naturalizada. En su lugar, se propone una *metafísica aplicada* que no se corresponde ni con la delimitación a priori de ontologías posibles (metafísica a priori), ni con su subordinación a los criterios teóricos, prácticos y metodológicos de las ciencias a la hora de ofrecer una imagen ontológica del mundo (metafísica naturalizada). Entendemos que la metafísica aplicada atiende tanto a los resultados de la metafísica como de la ciencia a la hora de ofrecer una caracterización ontológica del mundo más rica y adecuada. Es más: a partir de los casos particulares de la aptitud biológica y de los holobiontes se hacen evidentes dos características esenciales que resultan de la interacción entre metafísica y biología en este marco, a saber: la *fertilización mutua* y la *borrosidad.*

Los casos de fertilización mutua son aquellos en los que la biología contribuye a mejorar los conceptos y teorías metafísicas inicialmente empleados en filosofía de la biología para tratar un problema conceptual. El caso de la aptitud biológica ilustra esta relación de reciprocidad, dado que la consideración de algunas de las características propias de los fenómenos biológicos lleva a reconsiderar ciertos aspectos no triviales de la teoría disposicional de la causación. Es, como veremos, el caso de la existencia de funciones disposicionales que pueden disponer tanto a favor como en contra de la aptitud del organismo (los rasgos sexuales, por ejemplo), la transformación que experimenta la aptitud a lo largo de la vida del organismo, o el rol epistemológico que juegan los factores que intervienen y previenen la manifestación de esta disposición. Por su parte, los casos de borrosidad se producen cuando no es posible distinguir con claridad qué tipo de interacción está operando en un caso concreto. Es más: es posible que las dos formas de interacción estén presentes al abordar un determinado problema. En los casos de borrosidad, el tipo de interacción resultará más bien una cuestión de énfasis, que podrá residir en las teorías metafísicas o en las teorías, fenómenos y prácticas biológicas.

Por último, no quiero dejar de incluir en este resumen algunos de los problemas y cuestiones que han ido apareciendo a lo largo del desarrollo de esta tesis, y que constituyen futuras líneas de investigación en las que me gustaría continuar trabajando. En primer lugar, me gustaría profundizar en mayor medida en la noción de metafísica aplicada, su rol en la ciencia, y sus diferencias con respecto a las metafísicas a priori y naturalizada. En la actualidad encontramos distintas propuestas orientadas a ofrecer una caracterización de la metafísica distinta de la a priori y la naturalizada (French y McKenzie, 2012, 2015; Le Bihard y Barton, 2018). Considero que una mejor caracterización de la metafísica aplicada puede contribuir en esta dirección.

En segundo lugar, en el Capítulo 3 he utilizado conceptos propios del campo de la filosofía de la biología tales como ‘autonomía’, ‘agencia’, y ‘teleología natural’ para sustentar mis argumentos a favor de la naturaleza emergente de la aptitud biológica. Se trata de conceptos complejos y muy debatidos en la filosofía de la biología, que, en la medida en que mi aproximación ha sido principalmente metafísica, simplemente los he asumido y los he empleado sin profundizar en ellos. Una línea futura de investigación consiste en estudiar estos conceptos en mayor detalle a la luz de conceptos metafísicos relacionados.

Finalmente, en lo que respecta al debate sobre el estatus ontológico de los holobiontes, querría centrarme en mayor detalle en la caracterización de los holobiontes como individuos y, en particular, ofrecer una caracterización de los mismos como individuos caracterizados por propiedades emergentes. En lo que respecta a este punto, estoy trabajando ya, en colaboración con Javier Suárez, una propuesta preliminar de los holobiontes como entidades emergentes.

***

La metafísica de la biología es una nueva tendencia muy prometedora que sólo recientemente ha empezado a captar la atención de los filósofos de la biología. Un sector importante en filosofía de la biología considera actualmente la metafísica como una herramienta fructífera para abordar problemas conceptuales en biología. La interacción entre metafísica y biología está empezando a desarrollarse y tomar forma, por lo que es necesario más trabajo en este campo para conceptualizar mejor esta nueva tendencia. Confío en que esta tesis sirva para poder arrojar algo de luz en esta nueva dirección.

Table of Contents

Introduction

The architects of the Modern Synthesis tended to interpret the triumph of Darwinism as the result of the abandonment of metaphysics. According to authors such as Ernst Mayr, the history of pre-Darwinian biology was dominated by Aristotelian metaphysics, in particular, by essentialism and teleology. In this received view, the populational reinterpretation of evolution put an end to the essentialist understanding of the species concept, and natural selection allowed the reduction of final causes to mechanical ones. In the new philosophy of the Modern Synthesis, metaphysics was negatively linked to vitalism, that is, to the belief in a non-physical force that governs living beings as opposed to inert ones. And metaphysical interpretations of biological phenomena, insofar as they have recourse to essences and final causes, should be avoided.

Philosophy of biology, as a discipline, appeared during the 1970s as a philosophy of the Modern Synthesis. Philosophers of biology rejected metaphysics since they inherited the negative view advocated by the Modern Synthesis and reinforced by the logical positivism. Thus, they favored mechanistic and epistemological approaches to biology while rejecting metaphysical ones.

However, in the last few decades, metaphysics seems to have experienced a renaissance in the field of the philosophy of biology. When looking at books and articles in this discipline, it is not rare to find philosophers that take into account metaphysical theories and concepts when approaching conceptual issues in biology. Recent works, such as that of Daniel Nicholson and John Dupré on *Everything Flows: Towards a Processual Philosophy of Biology* (2018), are a good illustration of this tendency, which has also made some philosophers talk of a new Metaphysics of Biology (Guay and Pradeu 2017).

What is the reason for this recent revival of metaphysics? Why are philosophers again interested in metaphysics when approaching conceptual problems in biology? What is the reason for currently exploring the metaphysical implications of biology, as well as their specificity regarding the physical sciences?

Perhaps metaphysics was never really abandoned in the philosophy of biology. As some authors have argued, questions such as *what is a species?* or *what is a biological individual?* have always been metaphysical issues. Indeed, the received view concerning the role of metaphysics in philosophy of biology is changing among historians of both logical positivism and philosophy of biology. Michael Friedman has shown that the negative and naive view of metaphysics classically attributed to the logical positivists is not accurate (Friedman 1999). Concerning the history of the philosophy of biology, it has been recently argued that the metaphysical discussions on the specificity of living beings that took place between the two World Wars were not necessarily linked to vitalism (Nicholson and Gawne 2015). Moreover, recent empirical and conceptual advances in different biological sciences have made philosophers appeal to metaphysical concepts that allow for a better characterization of the complexity of biological phenomena.

This doctoral dissertation attends to this new emergent interaction between metaphysics and biology. Besides paying attention to the extant literature in the philosophy of biology in which this interaction takes place, I have myself metaphysically addressed two conceptual biological problems, namely that of fitness and that of the individuality of holobionts, which illustrate the interaction in a nice way. My metaphysical approach to these problems covers the largest part of this thesis. At the same time, they contribute to the characterization of the interaction between metaphysics and biology. In fact, as I will show, they instantiate two different ways metaphysics and biology interact, namely

Metaphysics *for* Biology (where metaphysics is used to clarify the reality biological concepts refer to) and Metaphysics *in* Biology (where the metaphysical implications of biological theories, phenomena, and practices are studied). Furthermore, in exploring how metaphysics and biology interact, I also pay attention to the kind of metaphysics that is playing that role: it is a form of metaphysics that does not correspond to the a priori or the naturalized ones widely acknowledged in the debate on the interaction between metaphysics and science in general.

The choice of the problems of fitness and the ontological status of holobionts is due to different reasons. Firstly, these debates are a good illustration of the usefulness of metaphysics in both classic and contemporary debates in philosophy of biology. Secondly, in the classic philosophical debate on the characterization of the concept of fitness, the necessity for an accurate ontological characterization of fitness as a property of organisms has been explicitly addressed, and metaphysical approaches to fitness have already been performed by philosophers of biology. Finally, concerning the ontological status of holobionts, metaphysical concepts such as that of 'individuals', 'persistence', 'composition', or 'parthood', do also permeate the debate.

The thesis is articulated in four chapters. In Chapter 1, I briefly introduce both the debate in metaphysics of science concerning the interaction between metaphysics and science and the current tendency in the philosophy of biology to approach conceptual problems in biology from a metaphysical perspective. These two topics constitute the conceptual framework of my thesis. It is from them that I approach the interaction between metaphysics and biology. In particular, I offer a categorization of two different ways metaphysics and biology interact (Metaphysics *for* Biology and Metaphysics *in* Biology, already mentioned above).

In the rest of the chapters, I show the results of my own analysis of conceptual problems in biology from a metaphysical perspective. In Chapter 2, I introduce the dispositional theory of causation and offer a characterization of fitness as a causal dispositional property of organisms. As I will argue, this account of fitness offers several advantages in solving the conceptual problems traditionally attributed to fitness in classic evolutionary biology.

In Chapter 3, I briefly introduce the metaphysical notion of emergence and the different forms of autonomy and dependence that are associated with it. I will explore what criteria of dependence and autonomy can be applied to fitness, and argue that fitness is a strong emergent property of organisms, insofar as it causally depends on its basis and it is causally autonomous. Finally, in Chapter 4, I pay attention to the debate on the ontological status of holobionts and I address the question of whether holobionts are individuals or ecological communities. In particular, I focus on the eco-immunity account of holobionts and explore the metaphysical implications regarding parthood that follow from its understanding of holobionts as hybrids between a host and its microbiota.

This thesis could have been written in different ways. Firstly, I could have approached the interaction between metaphysics and biology on the basis of different conceptual problems than those of fitness and holobionts. The choice I made, as I have said before, is due to the rich interaction between metaphysics and biology that I found in them. But I am aware that there are many other issues and concepts that are also good illustrations of such interaction. Secondly, in addressing each of these questions, I have paid attention to different metaphysical concepts and theories, such as that of dispositions, causation, dependence, emergence, parthood, or persistence, among others. But there are other interesting metaphysical tools that are useful and I have not paid attention to (the process/substance dichotomy is a case in point). Finally, even if Chapter 1 was the last chapter I wrote, I decided it to be the first one because, from a conceptual viewpoint, it introduces the general philosophical framework that encompasses the more specific goals achieved in chapters 2, 3, and 4. But of course, insofar as it presents my results concerning how metaphysics and biology interact, it might have also been the final chapter of this thesis. The conclusions include some reflections on the interaction between metaphysics and biology in relation to the cases of fitness and holobionts treated throughout the dissertation. In particular, I expose the reasons why the approach to the conceptual problem of fitness can be considered as a case of Metaphysics *for* Biology whereas the approach to the debate on holobionts is better conceived of as a case of Metaphysics *in* Biology.

The variety of topics I have dealt with in both metaphysics and philosophy of biology has forced me to limit the depth and detail with which I deploy the issues, so that clarity is guaranteed as much as possible. Here I also could have chosen just one of the issues and developed it in more detail to explore the interaction present in it. Yet, I decided to go over several issues to better illustrate the manifold and complex way in which metaphysics and biology interact.

This doctoral dissertation was initially motivated by a problem that has always puzzled me and that I have been able to conceptualize thanks to philosophy, namely what the fundamental structure of the world is and how we can obtain knowledge of the ontological status of the entities within it. Since I was a child, I have wondered about the change of things and the passage of time. My approach to philosophy and, in particular, to metaphysics, gave me the resources to initially address these questions.

At the beginning of my research, and due to my young inclinations, my sympathies were fundamentally metaphysical. In particular, I started getting used to metaphysical concepts such as that of causation, laws, persistence, or change. During my Degree in Philosophy, and thanks to María Cerezo, I had the opportunity to work on questions concerning persistence and change, and became familiar with the contemporary metaphysical theories of persistence, namely endurantism and perdurantism. My Master thesis focused on the consequences of the application of metaphysical theories of persistence to the case of biological species and, in particular, on the implications that such application has for the thesis of the metaphysical equivalence between three-dimensionalism and four-dimensionalism (Reydon 2008). Some of the results of this work were published (Triviño and Cerezo 2015). It was precisely in Reydon's work that I found a clear case of interaction between metaphysics and biology. In fact, he uses the term 'cross-fertilization' when talking about the influence that his conclusions in the philosophy of biology have for metaphysics. As I will argue, cross-fertilization is one of the features that characterize the interaction between metaphysics and biology.

During the first year of my PhD, my main interests remained metaphysical. In particular, I was interested in the characterization of dispositional and emergent properties. During this period, I was sponsored to attend the *Emergence Summer School,* at Durham University, wherein my interest in characterizing this kind of properties increased. My first approach to the issue of fitness was something casual. It took place when I read Alexander Rosenberg's paper on *The supervenience of the biological concepts* (1978), in which he characterizes fitness as a supervenient property. Since I did not fully agree on Rosenberg's characterization of fitness due to his understanding of supervenience, I started to look for a metaphysical framework that allowed for a better conceptualization of this biological notion. It was then when I first had to approach the literature in the philosophy of biology on fitness in more detail.

I had the opportunity to present an initial idea I was considering regarding fitness as an emergent dispositional property of organisms at the *3rd European*

Advanced Seminar in the Philosophy of the Life Sciences that took place at the Konrad Lorenz Institute for Evolution and Cognition Research (KLI) (Klosterneuburg, Austria). It was there where I met who would become the co-supervisor of this dissertation, Laura Nuño de la Rosa. She encouraged me to keep working on this line and helped me to improve the characterization of fitness as a disposition by guiding my philosophical readings on this topic. It was from that moment on that I started to consider the philosophy of biology in more serious detail besides metaphysics. At this point, it was crucial to read and study the dispositional theory of causation developed by Stephen Mumford and Rani Lill Anjum (2011), which provided the metaphysical tool I required to develop my project.

Thanks to a research stay at the KLI between March and July (2015), I had the opportunity to know more about some of the philosophical questions on other biological fields besides classic evolutionary biology, such as evolutionary developmental biology, or evo-devo. Professor Lee Altenberg helped me a lot in this regard. Furthermore, I also had the opportunity to talk to other philosophers of biology who were also working on issues related to fitness at that time. I am especially grateful to Lynn Chiu, who gave me very relevant feedback to improve the draft on the dispositional account of fitness I was working on.

A year later, between February and June (2016), I also enjoyed a research stay at the *Egenis the Centre for the Study of Life Sciences,* under the supervision of Professor John Dupré. It was there where I started to work in more detail on metaphysical notions such as that of emergence and supervenience and to develop further the arguments for the characterization of fitness as an emergent disposition. In this regard, Professor John Dupré helped me to improve this conceptualization of fitness by giving important feedback to the different drafts on which I worked there. A previous version of this work was presented at the *31st Boulder Conference on the History and Philosophy of Science. Topic: Emergence*

(Boulder, Colorado), where I met Professor Jessica Wilson, who discussed the project with me and encouraged me to keep working on this line. It was also during my stay at *Egenis* that I met Javier Suárez, with whom I had interesting conversations regarding emergentism, fitness, and holobionts. In fact, my philosophical approach to the problem of the ontological status of holobionts started after those conversations.

Although since I was a student in the Philosophy Degree my interests have always been metaphysical, my approach to some problems in the philosophy of biology has allowed me to realize of two important issues, namely that metaphysics can be a useful tool to address conceptual problems in biology; and that, contrary to what I initially thought, metaphysics, by itself, is not able to offer an ontological worldview, but it also needs to consider what the different scientific disciplines claim about the world, at least regarding their particular subject-matter. I worked in more detail on these questions concerning the connection between metaphysics and science during my research stay in Madrid (December-January 2017), wherein I participated in the activities organized by Professor Mauricio Suárez in the *Seminars on Methods of Causal Inference and Scientific Representation (MCISR).*

Some of the results in this thesis have been presented to the philosophical community in different conferences, workshops, and summer schools, both national and international, I have published some of the results (Triviño and Nuño de la Rosa 2016), and have other manuscripts under evaluation or about to be submitted. I am in debt with all the people I have met and discussed with, some of whom I have mentioned above, in particular with my supervisors. I also thank the different institutions that have offered financial support for the development of my thesis. I have received a four-year fellowship from the *Fundación Séneca: Agencia de Ciencia y Tecnologí*a (Government of Murcia, Spain), that has also

covered my research stays. Furthermore, my attendance to different conferences, workshops, and summer schools has also been possible due to the financial support of Maria Cerezo's projects on *Contemporary Issues in the Metaphysics of Biological Sciences (*Ref. 19489/PI/14), funded by "Fundación Séneca. Agencia de Ciencia y Tecnología de la Región de Murcia", and *Metaphysics of Biology: Framing the Interactions between Metaphysics and Molecular, Developmental and Evolutionary Biology* (Ref. FFI2017-87193-P), funded by the Ministry of Economy and Competitiveness (Spanish Government).

Chapter 1:
Towards a Metaphysics of Biology

> *"I am quite sure that many of the two hundred or so experimental papers I produced have been definitely affected by consciously held metaphysical beliefs, both in the types of problems I set myself and the manner in which I tried to solve them."*
> (Waddington 1969)

Since the origins of the philosophy of biology as a discipline, around the 1970s, philosophers have mostly approached theoretical and conceptual problems arising in the biological sciences from an epistemological and methodological perspective. However, since the end of the 20th and the beginning of the 21st century, philosophers of biology have tended to acknowledge the usefulness of metaphysics when dealing with this kind of problems (Boogerd et al. 2005; Huneman 2010; Dupré 2012, 2015; Austin 2016b, 2017; Waters 2017; Nicholson and Dupré 2018). This current interest in metaphysics indicates that a new form of performing philosophy of biology, one that has been named 'Metaphysics of Biology' (Guay and Pradeu 2017), is getting more common.

The increasing appeal to metaphysics by philosophers of biology is a phenomenon that deserves more attention in order to clarify how they use metaphysics when approaching conceptual problems in biology, i.e. how metaphysics and biology interact. My reflection on the interaction between metaphysics and biology has resulted from my analysis of extant literature in philosophy of biology, but has mainly taken shape in the course of my own work as a participant of this interaction. In particular, I have addressed two conceptual problems in biology from a metaphysical perspective in order to clarify them, namely the classic question in evolutionary biology concerning the characterization of fitness, and the more contemporary question concerning the ontological status of holobionts.

This twofold research has led me to realize that the interaction between metaphysics and biology seems to take place in more than one way. Moreover, I have concluded that the kind of metaphysics that is involved in these interactions does not seem to fit in the classic forms of interactions that are usually considered in the metaphysics of science literature. In this chapter, I have a twofold interrelated aim: on the one hand, I propose a categorization of the different forms of interaction between metaphysics and biology and, on the other, I try to clarify the kind of metaphysics that seems to be at play in each of these interactions.

The results of exploring both the kind of interaction between metaphysics and biology, and the form of metaphysics involved in this interaction, might allow us to shed some light on the characterization of metaphysics of biology, and more generally, on the relationship between metaphysics and science. Generally, the debate regarding the interaction between metaphysics and science has been framed by considering only one particular science: physics (Ladyman and Ross 2007, 2013; Ross et al. 2013). This is what has been referred to as 'The Priority of the Physics Constraints' (Melnyk 2013). Recently, notwithstanding, it has been argued that this approach offers a too simplistic view of the interaction between metaphysics and science and that other sciences besides physics should be also considered. As Alexandre Guay and Thomas Pradeu state it:

> The second oversimplification concerns an undue restriction of scientific metaphysics to the three theses mentioned above, namely reduction to physics, universalism, and realism. The idea that all reality "ultimately reduces" to physics is far from obvious. More importantly, it is simply not true that all projects in science-based metaphysics have been rooted in physics (even if one could argue that they have been assuming ontological physicalism) [...] It is also far from obvious that all projects in scientific metaphysics should profess exclusiveness, as if there was only one metaphysical approach or valid ontology. *Metaphysics is, in general, a uniforming enterprise, but this does not mean that one single approach will be the only appropriate one, at the expense of the others* [...] The view that one given approach is uniquely appropriate and fundamental (in the present case, a metaphysics derived from our best current physical theories) cannot be accepted without demonstration (Guay and Pradeu 2017, 7. Emphasis added).

The concern regarding the predominance of physics in the debate over metaphysics of science is similar to that concerning the characterization of the philosophy of science during the beginning of the 20th century. During this period, philosophy of science was mainly equated to philosophy of physics, and other scientific disciplines (including biology, anthropology, psychology or sociology) were also approached from the conceptual framework of the physical sciences (Mayr 2004, 16-17).

The equation between philosophy of science and philosophy of physics was driven by the consideration of physics as the fundamental science and the possibility, at least in principle, of reducing all the different special sciences to this fundamental one (Oppenheim and Putnam 1958). However, since the 1990s, an increasing number of biologists and philosophers started to vindicate the autonomy of biology from physics. Ernst Mayr was one of the major advocates of this thesis. In his view, the consideration of certain basic physicalists concepts that are not clearly directly applicable to biology (e.g. essentialism, determinism or reductionism), and the recognition of irreducible properties in the living entities due to their holism and the complexity of the interaction among their parts (e.g. metabolism, replication, adaptedness, growth, hierarchical organization), allow the justification of the autonomous character of biology (Mayr 1982, 2004).

A similar situation has taken place in the debate concerning the relationship between metaphysics and science. This debate has historically been established in terms of how metaphysics interacts with or plays a role in relation to physics. However, the acknowledgment of the interaction between metaphysics and biology, due to the increasing use of metaphysics by philosophers of biology, as well as the study of how this interaction is performed, could allow us to shed new light on this debate. In this sense, the conclusions obtained from the exploration of the interaction between metaphysics and biology can serve to enrich the more general debate regarding the interaction between metaphysics and science.

In this chapter, I will proceed as follows. Firstly, I introduce the debate in metaphysics of science concerning the interaction between metaphysics and science (§ 1). In doing so, I will pay attention to the two main positions that have characterized this debate, namely the view according to which metaphysics can be conceived of either as a guide or as a condition for the possibility of science (§ 1.1), and the view according to which science either guides or restricts metaphysics (§ 1.2). Secondly, I briefly introduce the recent history of the philosophy of biology, with an emphasis on the positivistic flavor that characterized this discipline in its origins (§ 2). In section 3, I address the central aims of this chapter: I present a categorization of different ways metaphysics and biology seem to interact, namely Metaphysics *for* Biology (§ 3.1) and Metaphysics *in* Biology (§ 3.2), while also exploring the kind of metaphysics that plays a role in each case. In section 4, I illustrate the complex character of the interaction between metaphysics and biology. Finally (§ 5), I conclude that Metaphysics *for* Biology and Metaphysics *in* Biology are instances of a current tendency in philosophy of biology, which I agree with philosophers of biology to label 'Metaphysics of Biology', and illustrate how the conclusions in this debate over the interaction between metaphysics and biology can serve to enrich the more general debate on the interaction between metaphysics and science.

1. Metaphysics of science

Philosophy of science, as an independent discipline, appeared during the 1920s due to the movement of logical empiricism. By the early 20th century, different sciences such as physics, mathematics, biology, chemistry, and the social sciences, were already independent of philosophy. In this context, the logical empiricists considered that the only way philosophy could continue being useful was by being scientific in character (Creath 2014).

Due to the influence of logical empiricism, philosophy of science focused on epistemological and methodological questions referring, mainly, to the nature of science and scientific criticism. Some of these questions are: 'what is scientific knowledge?', 'what is the scientific method?', 'what is the relation between scientific theories and truth?', 'what is scientific progress?', among others (Bird 1998b, 2011). Generally, logical empiricists approached this kind of questions on the basis of empirical observation, the inductivist logic, and verificationism (Bird 1998b). These methodological and epistemological approaches to science led to the rejection of metaphysics, understood as a discipline that can obtain an a priori, and, in some sense 'deeper', knowledge of the ontological structure and fundamental aspects of the world to which empirical sciences have no access (Creath 2014).[1] As Rudolf Carnap says:

> In the domain of metaphysics, including all philosophy of value and normative theory, logical analysis yields the negative result that all alleged statements in this domain are entirely meaningless. Therewith a radical elimination of metaphysics is attained, which was not yet possible from earlier antimetaphysical standpoints. It is true that related ideas may be found already in several earlier trains of thought [...] but it is only now when the development of logic during the recent decades provides us with a sufficiently sharp tool that the decisive step can be taken (Carnap 1959, 60-61).

In the 1960s, logical empiricism ceased to be cohesive enough to be identifiable as a movement, and by the 1970s it was completely over (Creath 2014; Uebel 2016). Two turns within the philosophy of science can be recognized from that moment (Papineau 1996; Bird 1998b). The first one took place in the 1970s, with the fall of logical empiricism, and it was driven by the acknowledgment of the historical, social, and also metaphysical character of science. During this period, authors such as Thomas Kuhn, Paul Feyerabend or Imre Lakatos, among others, challenged the view of science offered by logical empiricists, according to which an objective image of the world can be

---

[1] The logical empiricists' conception of metaphysics has its roots in the image of metaphysics presented by David Hume and Immanuel Kant. This image, notwithstanding, does not refer to the kind of metaphysics developed by the Greeks and the early scholastic (Boulter 2013). Furthermore, it has been recently argued that logical empiricists, in fact, endorse a kind of metaphysics. According to Michael Friedman, the traditional view of logical empiricists as committed to naive empiricism, and therefore, to a rejection of metaphysics, is not really accurate. In fact, logical empiricism, Friedman claims, is better understood as an intermediate position between the Kantian tradition and traditional empiricism (Friedman 1999).

obtained through the sciences, by criticizing the simplistic way of conceiving the role that experience and observation play in the development of scientific theories. Concerning the acknowledgment of metaphysics, in 1975, John Watkins claimed that the 'counter-revolution' against logical empiricism had triumphed: "I have the impression--he concluded--that it is now almost as widely agreed that metaphysical ideas are important in science as it is that mathematics is" (Watkins 1975, 91).

The second turn took place in the 1990s. It is characterized by both a comeback to the epistemological questions of science (although this time by taking into account the sociological and historical aspects of science itself) and a wider acknowledgment of the metaphysical implications and questions given in science (Papineau 1996; Bird 1998b; Gervais 2015). Due to this second turn, the philosophy of science was divided into two subdisciplines, namely the epistemology of science and the metaphysics of science (henceforth MOS) (Papineau 1996). According to David Papineau, the epistemology of science, which mainly concerns the justification of the scientific knowledge, is the branch that stands out within the philosophy of science. MOS, conversely, has less weight within the philosophy of science insofar as it is a more heterogeneous area wherein the problems that are part of it are more difficult to delimit.

In order to clarify the kind of problems that constitute MOS, Papineau distinguishes between two main groups of questions, namely general metaphysical questions and particular metaphysical questions. *General metaphysical questions* refer to the characterization of those metaphysical concepts that are common to all sciences, such as causation, laws or natural kinds (Papineau 1996; see also Bird 1998b). *Particular metaphysical questions* refer to the specific metaphysical questions that arise within a particular scientific discipline, such as the relative or absolute nature of space and time in physics (Papineau 1996). These two groups of questions have also been respectively referred to as the 'wide sense' and the 'narrow sense' of MOS (Hüttemann 2013).

The characterization of MOS as a subdiscipline within the philosophy of science presupposes a positive relation between metaphysics and science in which metaphysics, besides the methodological and epistemological approaches, is perceived as useful for philosophers of science. This positive relation presupposed between metaphysics and science was illustrated in two different ways. For some advocates of MOS, metaphysics plays the role of being a *guide* to science, that is, it is metaphysics that helps science to decide what problems need to be studied (Agassi 1975, 1996). Yet, other advocates have defended metaphysics to be not just a guide, but rather the *condition for the possibility* of science: metaphysics offers the conceptual framework that makes science possible (Wartofsky 1967).

Between the end of the 20th and the beginning of the 21st century, the conception of MOS changed. Instead of referring to a subdiscipline within the philosophy of science that explores the general and/or the particular metaphysical questions that appear regarding science, it started to refer to a discipline that studies the kind of interaction that is given between metaphysics and science (Esfeld 2007). In particular, it explores different alternatives and positions that can be defended concerning the role that metaphysics plays in science, as well as the role that science plays in metaphysics (Soto 2015, 26-27).

Currently, the debate in MOS concerning the interaction between metaphysics and science is generally divided into two main different positions. On the one hand, there has been a revival of the classic, positive perception of the relation between metaphysics and science referred above. In this sense, metaphysics is understood either as a guide to science or as a condition for the possibility of science itself. I will call this position 'the metaphysics prevalence view'. On the other hand, some authors argue that it is science that either guides or restricts metaphysics. I will refer to this position as 'the science prevalence view'. In the following sections, I will explain the metaphysics prevalence view (§ 1.1) and the science prevalence view (§ 1.2) in more detail.

1.1. The metaphysics prevalence view

According to the metaphysics prevalence view, metaphysics can be either a guide or a condition for the possibility of science. As noted above, this approach is a revival of the original characterization of MOS (Soto 2015). Currently, there are some advocates of this position, which has been mainly reduced to the idea that metaphysics is the condition for the possibility of science, what Soto has called 'maximalism' (Soto 2015). In this regard, Edward Jonathan Lowe has claimed that "[...] metaphysics goes deeper than any merely empirical science, even physics, because *it provides the very framework within which such sciences are conceived and related to one another*" (Lowe 2002, vi. Emphasis added).

Besides E.J. Lowe, Tuomas Tahko (2008), Anjan Chakravartty (2010, 2013a), or Stephen Mumford and Matthew Tugby (2013), are also some advocates of this approach. They offer different arguments to justify that it is metaphysics that makes science possible. According to Chakravartty, for instance, science relies on presuppositions that are beyond what is empirically accessible, and thus, it presupposes metaphysics (Chakravartty 2010, 70).

Furthermore, he claims, scientists are interested in studying the unobservable, i.e. those entities that are not directly empirically assessed but that serve to better explain the observable phenomena. And, in doing so, scientists turn to forms of speculation and methods that are not proper of science but of metaphysics (Chakravartty 2010, 63).

In a similar vein, Mumford and Tugby (2013) claim that metaphysics refers to the study of the fundamental aspects of reality, such as "kindhood, lawhood, causal power and causation", that make science possible insofar as they "impose order on the world" (Mumford and Tugby 2013). Other kinds of questions (e.g. time, space, composition, identity or persistence) could not be considered as part of MOS insofar as, they claim, do not refer to questions concerning the fundamental aspects of reality, but to particular metaphysical questions that appear within scientific disciplines such as physics.[2] In this sense, Mumford and Tugby reduce MOS to what Papineau characterized as general metaphysical questions (§ 1), and, based on this conception of metaphysics, they define MOS as follows:

> MOS def: The metaphysical study of the aspects of reality, such as kindhood, lawhood, causal power, and causation, which impose order on the world and make our scientific disciplines possible (that is, disciplines which are able to provide predictions -often novel- and offer explanations for new facts and anomalies within their given domain), and also the study of the metaphysical relationship between the various scientific disciplines (Mumford and Tugby 2013, 14).

Although the question concerning the kind of relationship that exists between metaphysics and science, and the question about how metaphysics

---

[2] As I will illustrate in the following section, this idea has been questioned by some advocates of the science prevalence view.

is conceived are different, they are obviously related. Thus, in the metaphysics prevalence view, metaphysics is generally characterized as 'a priori metaphysics' or 'metaphysics simpliciter' (Allen 2012; Chakravartty 2013a). In this sense, metaphysics is understood as a discipline that sheds light on the ontological structure of the world by means of a priori reflections that are made in order to establish what is possible (Chakravartty 2013a). The interaction between metaphysics and science, in this case, takes place insofar as metaphysics a priori delimits what the possible ontological features of the world could be, and it is science that establishes, in an empirical, and therefore, a posteriori way, what of those possibilities is the current, actual one (Lowe 2008; Tahko 2008, 2013).

The relevance of a priori metaphysics has been highlighted by authors such as Earl Conee and Theodore Sider (2005) or Tahko (2008, 2013), who characterize metaphysics as the science of the possible (see also Callender 2011). In this view, metaphysics is about "the most explanatory basic necessities and possibilities. Metaphysics is about what could be and must be. Except incidentally, metaphysics is not about explanatorily ultimate aspects of reality that are actual" (Conee and Sider 2005, 203). For the advocates of this position, metaphysics, insofar as delimiting what is possible, is ontologically and epistemologically prior with regard to natural sciences. As Tahko claims:

> We reach information about the (metaphysically) possible ways that the world might be with the help of a priori reasoning, which is ultimately grounded in essences, and the status of these results in terms of the actual world is determined by a posteriori [scientific] means (Tahko 2008, 266).

In the same line as Tahko, Chakravartty has also argued for the characterization of metaphysics as a science of the possible. He bases his arguments on the unobservables (Chakravartty 2010). In this regard, Chakravartty claims, the metaphysician may affirm the existence of unobservables such as quarks or possible worlds by means of a priori speculations. However, the scientist would remain agnostic about them and refrain from affirming or denying metaphysical claims regarding unobservables until she has empirical knowledge of them. According to Chakravartty, this illustrates the priority role of metaphysics with respect to science. The factual claims endorsed by scientists are a mere subset of those endorsed by metaphysicians insofar as metaphysicians would also endorse claims about unobservables whereas scientists would not directly do so (Chakravartty 2010, 69).[3]

1.2. The science prevalence view

The characterization of MOS from the perspective of the metaphysics prevalence view, according to which metaphysics is a condition for the possibility of science, has been challenged by those authors who defend what I term 'the science prevalence view' (Esfeld 2006, 2007, 2013; Ladyman and Ross 2007; Ross et al. 2013; Robus 2014; Soto 2015). The upholders of this view argue that it is only by following the scientific results that we can obtain an accurate ontological knowledge of reality. In this sense, they follow the line of Willard V. O. Quine in claiming that "it is within science itself, and not in

---

[3] I will come back to a different position and a criticism of Chakravartty's claim offered by Soto's 2015 in the next section.

some prior philosophy, that reality is to be identified and described" (Quine 1981, 21). Science is the only way to obtain knowledge about reality.

Two main approaches are generally recognized in the science prevalence view (Soto 2015), namely that according to which science is a *guide* to metaphysics, and that considering that metaphysics is *motivated* and *restricted* by the best current sciences. For the first approach, the results of our current sciences are what help us to answer metaphysical questions. In this sense, when there is a debate between two different metaphysical theories concerning a certain realm of reality, it is science, i.e. the results of a particular scientific discipline, that decides which of them is the correct one. The preferred metaphysical theory would be the one that has more theoretical proximity to and coherence with what our best available scientific theories claim about the world.

An illustration of this position is the metaphysical debate between presentism, i.e. in simple terms, the view that only the present is ontologically real; and eternalism, i.e. in simple terms, the view that past, present, and future are ontologically real. According to Sider, the special theory of relativity solves the metaphysical debate between presentism and eternalism by showing that the present has no distinctive ontological status, and therefore, that presentism is false. In this sense, Sider claims:

> In cases of science versus metaphysics, historically the smart money has been on science. At any rate, the present discussion will assume that consistency with something fairly close to current physics is a *constraint* that must be met by any adequate theory of time (Sider 2001, 42. Emphasis added).

Despite the subordination of metaphysics to science, for the advocates of this approach, there is still room for a form of metaphysics that is independent of science. Metaphysical theories concerning the ontology of the world can be proposed independently of science, but it is science that decides which metaphysical theory should be accepted and which one should not. This is the reason why science is said to guide metaphysics.

This position concerning metaphysics is different from that advocated by those who claim that metaphysics is *motivated* and *restricted* by the best current sciences. According to the advocates of this view (Ladyman and Ross 2007; Ross et al. 2013; Soto 2015), there is no room for a science-independent or a priori metaphysics insofar as this form of metaphysics cannot properly say anything about the structure of the world, not even can it establish what is possible. In this sense, "if metaphysics is to be part of the pursuit of objective knowledge, it must be integrated with science" (Ladyman and Ross 2013, 109). James Ladyman and Don Ross illustrate this position as follows:

> However [...] we deny that a priori inquiry can reveal what is metaphysically possible. Philosophers have often regarded as impossible states of affairs that science has come to entertain. For example, metaphysicians confidently pronounced that non-Euclidean geometry is impossible as a model of physical space, that it is impossible that there not be deterministic causation, that non-absolute time is impossible, and so on. Physicists learned to be comfortable with each of these ideas, along with others that confound the expectations of common sense more profoundly (Ladyman and Ross 2007, 16-17).

For the advocates of this position, therefore, metaphysics can be neither a condition for the possibility nor a guide to science. In this regard, Soto has recently claimed that Chakravartty's arguments are not enough to justify that science is metaphysically inevitable (Soto 2015). According to Soto, when scientists try to access and clarify unobservable entities such as electrons, the dark matter, or the Higgs boson, they do not turn to conceptual analysis or a priori metaphysical speculation, but to the development of experimental devices, the set of experiments, the construction and manipulation of models and theories, the use of data-gathering processes, and the application of mathematics, among others, that make those entities accessible (Soto 2015, 35).

In the same line, Soto also argues against Mumford and Tugby's (2013) characterization of metaphysics. On the one hand, he states that it is not clear that there is a fundamental and ordered structure of the world. This thesis, he claims, "is not straightforwardly suggested by science and requires, by contrast, a good deal of speculation which already finds its rationale in some form or another of metaphysics" (Soto 2015, 38). On the other hand, he also makes explicit the idea that concepts such as lawhood, kindhood, causation, and causal power do not seem to be the only concepts common to all sciences, but concepts such as *time* and *space* should be considered as well (Soto 2015, 38).

Besides the previous criticisms, the metaphysics prevalence view has also been criticized on the basis of the 'a priori' characterization of metaphysics it accepts. The arguments against a priori metaphysics can be generally classified into three types (French and McKenzie 2015). First, it is *frivolous*: insofar as it refers to mere possibilities, it ends up giving rise to strange thoughts such as those referring to possible worlds inhabited by zombies (see Chalmers 1996). Second, it appeals too much to *intuitions*: metaphysicians tend to appeal to their intuitions in order to justify their own theses and reject rival ones. And third, it is too *domesticated*: a priori metaphysics interprets questions and theories in contemporary physics as if they refer to more classic questions concerning fundamentality.[4] Thus, a priori metaphysics tends to present classic ideas and positions regarding ontology and fundamentality by using contemporary terms and debates in physics. Thus, it seems to properly match what current physics says about the world, although it is not really so, insofar as metaphysicians use these terms and theories as if quantum mechanics never happened (French and McKenzie 2015).

Concerning the science prevalence view, although arguing that science guides metaphysics and claiming that metaphysics is motivated and restricted by the current sciences might seem to be qualitatively different positions, both coincide on the idea that the only kind of proper metaphysics is the one that follows from the scientific view of reality and that is constrained by it. Thus, although upholders of the idea that science guides metaphysics consider that there is room for a science-independent metaphysics, ultimately, the only metaphysics they accept is the one that is more proximal, coherent, and therefore *constrained,* by the best current science (Sider 2001, 42).

This form of science-dependent metaphysics has been called 'naturalized metaphysics' (see, for instance, Ross et al. 2013; Esfeld 2018), and it is considered to be enough, in the sense that it is not necessary to recognize other forms of metaphysics besides this one (Soto 2015, 29). Naturalized metaphysics focuses on exploring the ontological worldview that *follows* from the different sciences and, in this sense, is the only one that properly interacts with science itself. In this view, therefore, naturalized metaphysics is the only

---

[4] I am aware that these criticisms to a priori metaphysics can be questioned as well. However, my intention in this section consists merely in presenting the different arguments and ideas proposed by the advocates of the science prevalence view.

one that is able to properly say something about the fundamental structure of the world.

In this regard, Guay and Pradeu have recently claimed that "[MOS] used to describe a philosophical project, which aims at developing an ontology or worldview based on current science" (Guay and Pradeu 2017). In the same line, John Dupré and Daniel Nicholson seem to be committed to naturalized metaphysics when claiming that "metaphysics is generally to be established through empirical means, and is ultimately therefore answerable to epistemology. Scientific and metaphysical conclusions do not differ in kind, or in the sort of arguments that can be given for them, but in their degree of generality and abstraction" (Dupré and Nicholson 2018, 4).

Ladyman and Ross (2007) illustrate their commitment to naturalized metaphysics by means of quantum theory. According to them, the kind of ontology or worldview that can be followed from quantum theory is one they refer to as 'ontic structural realism' (see also Ladyman 1998; French and Ladyman 2003a, b; and Esfeld 2004. See Chakravartty 2007 for some critiques to this position). According to ontic structural realism, structural relations are ontologically prior to the entities and properties that are given in the world. In fact, both entities and properties are determined by or depend on those relations:

> Ontic Structural Realism (OSR) is the view that the world has an objective modal structure that is ontologically fundamental, in the sense of not supervening in the intrinsic properties of a set of individuals. According to OSR, even the identity and individuality of objects depend on the relational structure of the world. Hence, a first approximation to our metaphysics is "There are no things. Structure is all there is" (Ladyman and Ross 2007, 130).

Insofar as the science prevalence view considers all the questions concerning the ontological status of the world to be solved by our best current sciences (Melnyk 2013), naturalized metaphysics has to perform its task in *continuity* with them (Ladyman and Ross 2007; Esfeld 2007; Ross et al. 2013). As Chakravartty states it, "continuity with science [...] is the suggested means by which to ensure that metaphysics does not lapse into the unprofitable excesses of non-naturalized metaphysics. Naturalized metaphysics, ex hypothesis, in virtue of its continuity with science, enjoys some degree of epistemic privilege" (Chakravartty 2013a, 40). Although the characterization of *being continuous* with current sciences is vague in itself, it is generally used to mean *being 'constrained'* by the current sciences (Allen 2012; Chakravartty 2013a). Here, the constraint refers to the idea that metaphysics has to be compatible with the methodological, epistemological, and ontological criteria that science imposes, in order to obtain an accurate knowledge about the structure of the world and to offer an ontological worldview (Allen 2012; Soto 2015). In other words, the scientific practice exemplifies and provisionally establishes epistemic, methodological and ontological criteria, which should work as desiderata in the pursuit of metaphysics, thereby constraining metaphysical practice (Soto 2015, 53).

Naturalized metaphysics, notwithstanding, has also received some criticisms. One of them is based on the idea that, if metaphysics is constrained by scientific epistemological, methodological, and ontological criteria, then it is not able to perform the task it is supposed to do. As Chakravartty has made explicit, the difference between a priori and naturalized metaphysics does not lie in the kind of task they perform, that is, to shed light on the ontological features of the world. Rather, the difference between these two kinds of metaphysics consists in the way they perform their tasks, that is, by means of a priori stipulations about what is possible, or by means of being constrained by the best current sciences, respectively (Chakravartty 2013a, 32).

In this line, Sophie Allen has recently argued that naturalized metaphysics is not properly a form of metaphysics insofar as it does not really perform the task it is supposed to do: it does not shed light on the ontological features of the world or offer an accurate ontological worldview (Allen 2012). The reason, Allen argues, is that naturalized metaphysics is compatible with ontological plurality, i.e. with the idea that different ontological worldviews might follow from a particular scientific theory, and it does not have the resources to compare and select among them.[5]

Allen illustrates this limitation of naturalized metaphysics by means of an example (Allen 2012, 219). She argues that, in what concerns ontic structural realism, there are two possible ontological positions that are compatible with it, namely (1) *the non-eliminativist view*, according to which entities and their properties have a derivative existence, i.e. they are not fundamental but they exist insofar as grounded on the causal and structural relations that are fundamental. And (2), *the eliminativist view*, which entirely rejects the existence of objects and properties. Ontic structural realism is compatible with ontological plurality in this sense. In terms of naturalized metaphysics, therefore, both positions are ontologically equal. It is not possible to choose one position as the correct ontological worldview insofar as both follow the methodological, epistemological or ontological criteria of science. In this regard, Allen argues, naturalized metaphysics is not able to perform its task of clarifying the ontological worldview. Instead, it seems that,

---

[5] There is a different form of conceiving ontological plurality, namely as the metaphysical thesis according to which the ontological worldview obtained from different scientific fields, theories, and practices, such as those of physics and biology, for instance, might be equally valid. This metaphysical thesis is the one I accept and presuppose in this thesis.

in order to answer this kind of questions, we need to appeal to a form of 'a priori' metaphysics, insofar as "in 'a priori' metaphysics, the ontological alternatives are clearly different theories [...] [while] naturalized metaphysics leav[es] it open to question whether there is any difference between these theories at all" (Allen 2012, 230). However, appealing to a priori metaphysics is not really an option for the naturalized metaphysician since this form of reasoning is beyond the methodological, epistemological, and ontological criteria of science. As Allen states it:

> [T]he methods which are acceptable within naturalized metaphysics are not sufficient to make a choice [between different ontological theories], and the methods which do allow one to make a choice between ontological theories are not acceptable within the methodology of naturalized metaphysics. Thus, the naturalized metaphysician lacks the resources with which to determine (or even approximate) which ontological theory is the correct one [...] (Allen 2012, 221).

The advocates of naturalized metaphysics might claim that these kinds of questions concerning the clarification of the accurate ontological worldview are beyond their task. Our scientific theories might be compatible with different ontological worldviews but it is not a task for the naturalized metaphysician to choose among them. In my view, this is not an accurate answer, notwithstanding. Following Chakravartty's characterization of metaphysics and Allen's argument, I consider that, insofar as it is a form of metaphysics, naturalized metaphysics needs to be able to clarify how the world ontologically is and to offer arguments for favoring one ontological worldview over another. If not, then we do not have a form of metaphysics but something different.

Besides its impossibility to avoid ontological plurality in Allen's sense, naturalized metaphysics has also been challenged on the basis that it already assumes a priori metaphysical commitments (French and McKenzie 2012, 2015; Chakravartty 2013a). French and McKenzie (2012, 2015) have illustrated this position by means of the ontic structural realism proposed by Ladyman and Ross. In ontic structural realism, they claim, the fundamentality of the structural relations already presupposes metaphysical commitments concerning the characterization of the fundamental, the structure, or the relation between the fundamental and the non-fundamental. These presuppositions, however, are not explicitly recognized, and they cannot even be clarified by naturalized metaphysics, but rather by another form of metaphysics that might be in no way related to science (French and McKenzie 2015).

***

As illustrated in sections 1.1 and 1.2, in both the metaphysics prevalence view and the science prevalence view, the debate in MOS concerning the interaction between metaphysics and science is mainly approached from the frame of physics (Ladyman and Ross 2007, 2013; Ross et al. 2013). This frame, notwithstanding, could be extended if other scientific disciplines are taken into account. In this sense, the recent increase of the interaction between metaphysics and biology among philosophers of biology might serve as a different frame from which to shed light on the MOS debate. As I see it, the consideration of the way philosophers of biology use metaphysics when approaching conceptual problems in biology can serve us to enrich the more general debate concerning both the interaction between metaphysics and science as well as the characterization of the form of metaphysics that plays a role in this interaction.

In this chapter, I offer a categorization of different forms in which metaphysics and biology seem to interact on the basis of a previous study of particular cases in which this kind of interaction takes place (Chapter 2, Chapter 3, and Chapter 4). As I see it, approaching the debate in MOS from the frame of the biological sciences evidences that the interaction between metaphysics and science is more complex than usually considered, i.e. as a dichotomy between the metaphysics prevalence view and the science prevalence view, and that the form of metaphysics interacting with the sciences do not seem to be only the a priori nor the naturalized one. Before exploring in more detail the interaction between metaphysics and biology (§ 3 and § 4) and what insights it can offer concerning the MOS debate (§ 5), in the next section I will offer a brief introduction to the recent history of the philosophy of biology and the way it has evolved in what regards the interaction between metaphysics and biology.

2. Philosophy of biology and the neglect of metaphysics

Although nowadays it has become a common procedure for some philosophers of biology to take into account metaphysics when addressing conceptual problems in the life sciences, this was not the way philosophers of biology tended to proceed in the past. Historically, the study of themes at the intersection between metaphysics and biology has been less common than those at the intersection between metaphysics and physics. Generally, this neglect of metaphysics has been attributed to the negative image that certain philosophers of biology had of metaphysics due to the influence of logical empiricism (Nicholson and Gawne 2015), as well as to the belief in the specificity of biology as opposed to the generality of physics. As Kenneth Waters states, "scientific metaphysicians assume that whatever is true of fundamental physics must be true of everything (and whatever is true of biology is not true of everything)" (Waters 2017, 85).

Although philosophers (e.g. Aristotle, René Descartes, Gottfried Leibniz, and Immanuel Kant) have long been interested in biology, philosophy of biology, as a discipline, is considered to have emerged in the 1970s. According to what can be called 'the standard view' among philosophers of biology, it was the fall of logical empiricism that gave rise to the philosophy of biology (Rosenberg 1985; Ruse 2000; Hull 1969, 2002; Callebaut 2005). After the demise of logical empiricism, the acknowledgment of the particularity of different scientific disciplines, such as biology and psychology, took place. It is in this context that philosophers started to pay attention to the biological sciences and to be aware of the internal conceptual problems that characterize the theories developed in this field (Callebaut 1993, 74).

For the standard view, before the 1970s philosophers did not properly pay attention to what was going on in biology, and the communication between philosophers and biologists was almost nonexistent (Hull 1969). In this regard, Ralph S. Lillie's review on James Johnstones' *Philosophy of Biology* (1914) is a good illustration of the philosophers' lack of interest and knowledge of the biological sciences:

> We have noted some errors in matters of biological detail [...]. The brief account of certain physiological processes seems somewhat out of date; the account of the nerve impulse is unsatisfactory, and certainly few physiologists now hold that a muscle is a thermodynamic machine in the sense conceived by Engelmann; there is some evidence of unfamiliarity with biochemistry: the term "animo-acid" instead of amino-acid recur a number of times (Lillie 1914, 846).

The advocates of the standard view consider that the lack of communication between philosophers and biologists was, at least in part, a consequence of the philosophers' use of formalistic language. Due to the prevalence of logical positivism, the formalization of biological theories was the main way philosophers used to approach conceptual problems in biology. This formalistic language, besides being unintelligible to biologists, also deviated the attention from the biological problem at stake, since philosophers were more worried about whether a particular formalization was correct and whether there could be alternatives to it (Hull 1969).

In this sense, it was the fall of logical positivism, during the 1960s and 1970s, what gave rise to a new approach to biology, one in which biology itself was seriously considered and studied by philosophers. It is this new way of paying attention to biology by philosophers that is considered to be properly called 'Philosophy of Biology', or at least what Werner Callebaut called the 'New Philosophy of Biology' (Callebaut 2005, 99).

Recently, notwithstanding, the standard view has been challenged on the basis of the claim that there was a proper philosophy of biology before the 1970s, particularly, during the decades between the two World Wars (Bayron 2007; Nicholson and Gawne 2015). In this view, the philosophy of biology was already clearly delimited as a field of research during the first half of the 20th century due to the development of the organicist school. Organicists already dealt with biologists' work by directly reading and discussing their theories (Nicholson and Gawne 2015, 347-348). Before the 1970s, therefore, communication between philosophers and biologists was already taking place and the work of philosophers in relation to biology dealt with other issues than the mere formalization of theories.

While the standard view should be nuanced, the recognition that the fall of logical empiricism gave rise to a new approach to biology, one in which the subject of biology (discoveries, issues, advances, theories, methodologies…) was studied by philosophers such that the communication between philosophers and biologists increased, still holds. As stated above, it was since the 1970s that philosophers started to pay more attention to the biological sciences (particularly, to evolutionary biology), and, as a result, biologists started to consider the philosophers' work in a more serious way (Hull 2002; Callebaut 2005). In Callebaut's words:

> An important beneficial effect (although possibly unintended and unanticipated) of these shifts in philosophical style and focus is that after several decades of splendid isolation, philosophers are again talking to "their" scientists, and more and more working biologists develop an interest in what philosophy of biology -and in particular, philosophy of evolutionary biology- has to offer. One good indicator of this new dialogue is the rapidly increasing number of joint publications by biologists and philosophers [...]. Sober's *Conceptual Issues in Evolutionary Biology* (1994) is used in evolution courses for evolutionary biologists as well as for philosophers, often in conjunction with one of the textbooks that have become available in philosophy of biology (Callebaut 2005, 107).

According to Callebaut, philosophy of biology is the discipline in which philosophy contributes to improving our understanding of the relevant conceptual biological issues at stake (Callebaut 2005, 101). In the beginning, these conceptual issues were mainly those generated within the theoretical framework of the Modern Synthesis, such as those concerning fitness, natural selection, genes, or mutation, among others. However, since the late 1990s, and due to the new empirical and theoretical developments experienced by evolutionary biology, philosophers of biology started to pay attention to other conceptual issues beyond those that had worried the architects of the Modern Synthesis, such as those concerning evolvability, phenotypic plasticity, or epigenetic inheritance. In fact, a core epistemological debate among current theoretical biologists and philosophers of biology concerns how this new conceptual development should be integrated into evolutionary theory, some arguing that it constitutes an extension of the Evolutionary Synthesis (Pigliucci 2007; Pigliucci and Müller 2010).

Although for the standard view it was the fall of logical empiricism that gave rise to the philosophy of biology, this discipline continued having a positivistic flavor, as it is illustrated by the efforts made to axiomatize and formalize biological theories, such as the evolutionary theory (Williams 1970; Williams and Rosenberg 1985). Furthermore, different biological concepts, such as 'biological function' (Wright 1973, 1976; Cummins 1975; Boorse 1976, 2002; Bigelow and Pargetter 1987; Millikan 1989; Godfrey-Smith 1994; Mossio et al. 2009), 'information' (Godfrey-Smith 2000, 2008; Griffiths 2001; Shea 2011, 2013; Lean 2014), or 'adaptation' (Sober 1993; Lewens 2007), as well as questions, such as 'can evolution by natural selection and evolution by drift be differentiated?' (Millstein 2002, 2005; Brandon 2005), or 'can Mendelian genes be reduced to molecular genes?' (Schaffner 1969; Waters 1994), were mainly approached from an epistemological perspective.

In this regard, it is important to clarify that these questions can be approached from a metaphysical perspective as well. Waters, for instance, offers a list of different questions (e.g. 'what is a gene?', 'what is a mechanism?', 'what is a species?) that some philosophers of biology have addressed with the goal of identifying the fundamental aspect of reality they refer to (Waters 2017). While I agree with Waters that there are indeed some metaphysical approaches to concepts and questions in the philosophy of biology, my aim at this point is to highlight that this was not the main tendency among philosophers of biology. In this sense, I agree with Mayr's (2004) and Callebaut's (2005) description of mainstream philosophy of biology:

> [...] philosophers of biology, such as Bechtel, Hull, Kitcher, Rosenberg, or Sober, insist that the *high standards of clarity and rigor* introduced by logical empiricism deserve further emulation. Most of the philosophers of biology of the founding generation had their roots in analytic philosophy [...], if not directly in the logical-empiricist tradition (Callebaut 2005, 107).

In the philosophy of biology of this period, epistemological approaches to conceptual problems in biology were so common that, in some cases, it is difficult to properly distinguish philosophy of biology from theoretical biology. Indeed, some authors claim that there is no difference between them (Hull 2002), whereas others, such as Paul Griffiths, see it as a question of focus: philosophers approach biological problems from different perspectives, and they look for different kinds of answers when compared to theoretical biologists (Griffiths 2008).

Nonetheless, as previously noticed, some metaphysical questions did find a place in the philosophy of biology. The main, classic one is the question on the ontological status of species in evolutionary biology (Stamos 2003; Richards 2010; Kunz 2012), a question that, in fact, refers to two sub-questions, namely the *realism species issue* and the *species ontology issue*. The *realism species issue* concerns whether species are real entities existing in nature or just mere human constructions (Dobzhansky 1935; Ereshefsky 1998). The *species ontology issue* presupposes the reality of species, and focuses on the question about the kind of entities species are, e.g. classes (Hull 1965a) or individuals (Ghiselin 1971, 1974, 1997, 2007; Hull 1976, 1978, 1980).

Since the beginning of the 21st century, notwithstanding, the interaction between metaphysics and biology has become more common, and it does not only take place in what concerns classic evolutionary biology, but also in other biological disciplines such as evolutionary developmental biology (evo-devo), developmental biology, and molecular biology.

3. Interactions between metaphysics and biology: a categorization

The recent increase in the interaction between metaphysics and biology is a phenomenon that deserves more philosophical attention. The study of this interaction is not only relevant insofar as it might contribute to our understanding and characterization of a different tendency within the philosophy of biology, namely one in which philosophers of biology appeal to metaphysics when approaching conceptual problems in biology (Guay and Pradeu 2017; Waters 2017), but also insofar as it can help us to shed new light on the more general debate in MOS concerning the interaction between metaphysics and science as well as the kind of metaphysics that plays a role in it (§ 1).

In this section, I propose a categorization of two different forms in which metaphysics and biology interact. This categorization is aimed at illustrating the different ways *philosophers of biology* seem to appeal to metaphysics when approaching theoretical and conceptual problems in biology. However, during my study of the interactions between metaphysics and biology, I have also identified a different form of interaction in which it is the *metaphysician* who makes reference to different properties that characterize living entities in order to approach certain metaphysical questions and concepts. I will term this form of interaction, Metaphysics *with* Biology. To acknowledge this kind of cases is important since metaphysical questions, such as 'how do things persist through time?', 'what is composition?', 'when can we say that something has proper parts?', among others, are generally addressed by paying attention to inert objects and their properties. The general resource to statues and lumps of clay or gold to talk about persistence, identity, or the possibility of coincident objects are examples of this tendency (see Lowe 2002; Miller 2010).

In contrast, some metaphysicians have approached this kind of questions by taking into account specific features that characterize living beings and that are not present in inert entities. Peter van Inwagen (1990), for instance, recognizes the singular aspects of living entities when answering metaphysical questions about persistence, composition, and identity. Thus, with respect to composition, he claims that "(there is a y such as x_s compose y) if the activity of the x_s *constitute a life* (or there is only one of the x_s)" (van Inwagen 1990, 82. Emphasis added). Although there are problems regarding van Inwagen's understanding of 'life' (see Wartick 2012), it is the consideration of the living world and its proper features that allow him to establish his controversial conclusion (Hirsch 1993). In what he calls 'the Denial', van Inwagen claims that there are no tables, chairs or visible *objects* other than living organisms (van Inwagen 1990, 1. Emphasis added). According to his position, living organisms are the only visible material objects that exist since they are the only entities that are properly composed, i.e. the only entities whose components constitute a life. Other entities, such as tables or chairs, are not material objects, but rather a mere succession of aggregates.

This example illustrates that the consideration of features that are believed to be proper of living beings can influence the way a metaphysical question and/or concept is addressed and defined. In this regard, van Inwagen would not have claimed that the only visible material objects are living organisms if he had not recognized in them some form of composition and persistence that is not present in the inert world. In this form of interaction, nonetheless, the metaphysician's acknowledgment of what is proper of living entities is mainly phenomenological or, in some sense, *intuitive*. Metaphysicians do not generally appeal to *current* biological theories in order to develop their philosophical work. Yet, in spite of the phenomenological character of this approach, since (unlike the general tendency of metaphysicians to merely appeal to inert entities) the features of living entities are taken into consideration in these cases, and since their arguments do not contradict empirical science, I consider that this kind of cases can be considered as illustrating an interaction between metaphysics and biology.

Despite its importance, I will not dwell in more detail in this kind of interaction in the following. As previously noticed, my concern is how *philosophers of biology* can benefit from metaphysics when approaching theoretical and conceptual problems that arise in the biological sciences. In this sense, my main aim is at characterizing a current tendency within philosophy of biology, and thus, I will only pay attention to the way philosophers of biology appeal to metaphysics when performing their task. In this regard, I have recognized two different forms in which metaphysics and biology seem to interact: 'Metaphysics *for* Biology' (§ 3.1) and 'Metaphysics *in* Biology' (§ 3.2) (see Table 1).

	RELATION	FEATURES	EXAMPLES
METAPHYSICS *FOR* BIOLOGY	To use metaphysical concepts and/or theories in order to elucidate the aspects of reality to which biological concepts refer to as well as the features that that reality has.	-Metaphysics is used as a tool to make precise the aspect of reality to which a biological concept refers as well as to make precise its features. -Focus on metaphysical theories and concepts.	The ontological characterization of fitness (Rosenberg 1978, Brandon 1978, Mills and Beatty 1979).
METAPHYSICS *IN* BIOLOGY	Metaphysics studies ontological implications that are derived from biological theories, phenomena, and practices.	-Metaphysics is used as a tool to derive and study ontological implications of biological theories, phenomena, and practices. -Focus on biological theories, phenomena, and practices.	The debate over the existence or not of the final causes in nature (Mayr 1982).

Table 1. Summary of the preliminary taxonomy of the different ways metaphysics and biology interact.

3.1. Metaphysics *for* Biology

Philosophers of biology have generally approached biological sciences in order to clarify conceptual issues at stake from an epistemological and methodological framework (§ 2). However, it is currently common to find in the literature a metaphysical approach to this kind of problems as well. In these cases, philosophers of biology use metaphysical theories and/or concepts in order to elucidate the kind of ontological reality biological concepts refer to. In doing so, the conceptual issue at stake in biology is clarified. In other words, the philosopher is able to shed light on a conceptual problem in biology by means of clarifying (throughout metaphysical theories and concepts) the ontological status of the entity that biological concept refers to. I call this form of interaction 'Metaphysics *for* Biology' (Figure 1).

Some examples in the literature might serve to illustrate this current procedure within the philosophy of biology. Some of them include the use of the dispositional theory of causation to clarify the biological concept of the gene (Mumford and Anjum 2011, ch.10; Cerezo manuscript), the appeal to the metaphysical concept of emergence in order to characterize developmental modules in evo-devo (Boogerd et al. 2005; Huneman 2010; Nathan 2012; Walsh 2013; Brigandt 2015; Austin 2017), or the conceptualization of homologies and body plans as natural kinds (Hall 1996; Wagner 1996; Griffiths 1999; Rieppel 2005, 2006).

Metaphysics *for* Biology, therefore, pays attention to biological concepts, such as that of 'fitness', 'developmental modules' or 'homologies', among others. Metaphysical theories and concepts are used as a tool to elucidate the ontological status of the entities biological concepts refer to. In this sense, it seems that the kind of metaphysics that is performed in this case is not the a priori one. Metaphysics does not refer to the activity of delimiting what is ontologically possible concerning biological concepts such as the concept of fitness or the notion of gene. Philosophers of biology do not proceed independently of what biologists have said about the concept of concern, nor do they proceed without paying attention to the current biological theories. Instead, philosophers of biology take into consideration what biologists claim about the conceptual issue at stake and try to shed light on it in accordance with what biologists say. Thus, since philosophers of biology do not use metaphysics as a way of thinking of ontological possibilities without taking into account the biological sciences, the form of metaphysics they use is not the a priori one.

However, the form of metaphysics that seems to play a role in this case does not seem to be a naturalized one either. When exploring the ontological reality biological concepts refer to, these philosophers do not directly focus on obtaining the ontological consequences that follow from biological theories and practices. Rather, they use already given metaphysical theories and concepts, such as fundamentality, dependence, substances, dispositions or supervenience, and then apply them for the clarification of the conceptual biological problem at stake.

The problem in philosophy of biology concerning the ontological characterization of fitness might illustrate this point. In this regard, different metaphysical issues arise about fitness: how it is related to other properties (the view of fitness as a supervenient property), and what its nature as a property is (fitness as a disposition). The concepts of supervenience and disposition are metaphysical concepts widely discussed by metaphysicians. Philosophers of biology such as Alexander Rosenberg, Susan Mills and John Beatty, appeal to these notions, respectively, in order to clarify the ontological characterization of fitness. In the same line, as we will see in detail in Chapter 2, I do something similar when appealing to the metaphysical theory of causal dispositionalism proposed by Stephen Mumford and Rani Lill Anjum (2011), in order to clarify the ontology of fitness.

The metaphysical concepts and theories used by the philosophers of biology are part of the metaphysical toolbox (French and McKenzie 2012, 2015; French 2014), i.e. they are concepts and theories on which metaphysicians work without necessarily considering biology or any other science. In this case, therefore, it is the philosopher of biology who considers a particular metaphysical concept or theory as useful for approaching a conceptual issue in biology. In using metaphysical concepts or theories in order to address conceptual issues in biology, the philosopher of biology performs a kind of metaphysics that is not a priori nor naturalized, but a different one that I will term here *applied metaphysics*.

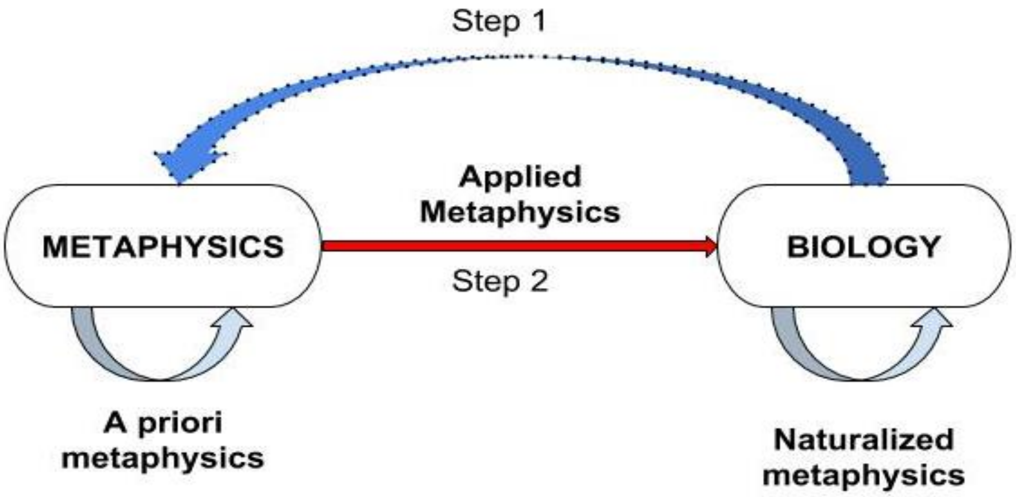

Figure 1. Metaphysics *for* Biology. The dotted dark blue arrow illustrates the first, previous step for the interaction between metaphysics and biology in this case. In it, the philosopher of biology is concerned with a conceptual problem in biology that she wants to address from a metaphysical perspective, i.e. by considering metaphysical theories and concepts. Step two is illustrated by the red arrow, and it refers to the step in which the interaction between metaphysics and biology is properly given. In particular, it refers to the actual application of particular metaphysical theories and/or concepts in order to clarify the ontological status of the biological concept at stake. In doing so, the conceptual issue in biology is clarified. In step 2, the kind of metaphysics the philosopher of biology is performing is the one I call 'applied metaphysics'. It is not the a priori nor the naturalized one, represented by the light blue arrows. In the first case, metaphysics will not take into account the results of biological sciences in order to establish the ontological reality biological concepts refer to. In the second, the ontological status of biological concepts will be established by merely following the epistemological, methodological, and ontological constraints of biological sciences themselves.

3.2. Metaphysics *in* Biology

Besides the previous form of interaction, there seems to be a different way in which philosophers of biology pay attention to the metaphysical aspects of biology. In this case, philosophers of biology do not directly appeal to metaphysical concepts and theories in order to address a conceptual issue in biology. Instead, they proceed by paying attention to biological *theories, phenomena,* and *practices*, in order to explore both the metaphysical implications they assume and the metaphysical consequences that might follow from them. And, in doing so, the philosopher of biology helps to shed light on the biological theories, phenomena, and practices themselves. I term this form of interaction Metaphysics *in* Biology (Figure 2). Different examples in the literature can also serve as illustrations of this form of interaction.

Concerning the ontological implications and consequences given in *biological theories*, Darwin's theory of evolution by natural selection has been a major focus of interest among philosophers of biology. Some of the possible ontological implications and consequences of this theory explored in the literature refer to the question about the existence of final causes in biology (Mayr 1982), the deterministic or stochastic status of evolution by natural selection (Horan 1994; Rosenberg 1994; Brandon and Carson 1996; Graves, Horan and Rosenberg 1999; Weber 2001; Bouchard and Rosenberg 2004), or the ontological status of species as individuals instead of classes (Ghiselin 1974, Hull 1978).[6] In this thesis, I also approach the metaphysical implications

---

[6] Regarding this debate, different ontologies have also been proposed besides that of classes and individuals, such as species as sets (Kitcher 1984a) -in which species might be both spatiotemporally restricted set of organisms, i.e. individuals; or spatiotemporally unrestricted sets of organisms, i.e. groups of organisms that share similar properties such as genetic, chromosomal and developmental; or species as homeostatic property cluster kinds (Boyd 1999a, 1999b; Griffiths 1999; Wilson 1999; Millikan 1999) -in which species are groups of

and consequences that follow from a biological theory. In particular, in Chapter 4, I pay attention to the eco-immunity account of holobionts (Chiu and Eberl 2016) and its characterization of holobionts as hybrids. As I will argue, this theory implies a particular form of conceiving parthood that gives rise to some problematic metaphysical consequences.

Philosophers of biology have also paid attention to metaphysical implications and consequences derived from *biological phenomena*. Thus, the issue of whether biological entities are processes or substances has become a core concern among philosophers of biology in the last few years. Although substance ontology, according to which substances are the primary units of reality, has been the predominant one in Western philosophy (Seibt 2016), advances in evolutionary developmental biology (Bapteste and Dupré 2013; Nuño de la Rosa 2013; Austin 2016b), as well as the study of some aspects of biological entities, such as developmental plasticity, robustness, or the different forms and properties that characterize organisms during the different stages of their life cycle, have recently been used to claim that biological entities are not ontologically substances but processes (Dupré 2012, 2015; Nicholson and Dupré 2018), that is, dynamic entities that are continuously changing (Seibt 2016). As Eric Bapteste and John Dupré put it:

> [A]s [evo-devo] has emphasized, an organism is a developmental process…[and] as evolution is uncontroversially a process, an evolutionary ontology will quite naturally be processual…[An] evolutionary ontology of the living world should distinguish the real evolutionary players, the units with causal powers resulting from or contributing to evolutionary processes (Bapteste and Dupré 2013, 380-381).

---

entities that share common properties or stable similarities that are not, in fact, essential for the membership in that species.

In particular, Dupré argues that biological entities are better understood as "a hierarchy of deeply intertwined processes, processes that are shaped by both higher and lower level processes with which they are connected" (Dupré 2015, 34). This metaphysical position of biological entities as processes has recently been challenged (Austin 2016b; Wiggins 2016). In this regard, Christopher Austin (2016b) claims that the biological aspects that are used in order to argue for a process ontology can also be explained in terms of a substance ontology:

> It's clear that although a process ontology undoubtedly represents a picture of the world which is scientifically unorthodox, it isn't empirically motivated –it may be theoretically costly, in that its framework demands that our fundamental categories be substantially restructured, but perhaps, if we wish to truly understand the empirical data of contemporary biological science, that's simply the cost of admission. […] On my view, adopting a process ontology amounts to buying into an inflated market –the theoretical benefits, I argue, have been overvalued. […] What I do want to argue, however, is that these are benefits that our current mechanistic [substance] ontology already affords us, and at no extra charge (Austin 2016b, 650).

Besides biological theories and phenomena, philosophers of biology also pay attention to the biological practice. Currently, there is a shift from a theory-based to a *practice-based metaphysics*. This new approach advocates the idea that it is possible to get some access to the structure of reality by focusing on scientific practices (Waters 2017). In this view, it is assumed that scientific practice develops as it does due to the fact that reality imposes some restrictions on it. Although this practice-based metaphysics has been generally developed in relation to physics, some examples of this kind of approach in biology can be found as well. Robert Brandon, for example, pays attention to the practice of evolutionary biology and concludes that there are no 'lawlike regularities' in this field, but only contingent ones. In this regard, he claims that:

> We can best make sense of the practice of experimental evolutionary biology if we see it as investigating contingent, rather than lawlike regularities. [...] I will argue that biologists are interested in contingent regularities, not for some purely sociological reason, but because the nature of evolutionary process. [...] experimental biology has the character it has because evolution produces contingent regularities (Brandon 1996, S444-S445).

In the same line, Waters has recently appealed to the use of the concept of 'gene' in contemporary genetics in order to explore the ontological implications that can be derived from it. In particular, he concludes that:

> Examining how the [molecular gene] concept is employed in practice reveals that it is not a category of being that "cuts nature at its joints". The problem with using this metaphor is that DNA has too many joints. [...] *the practice of genetics utilizes a flexible gene concept because there is no overall structure, functional or causal, of the elements of DNA* (Waters 2017, 83. Emphasis added).

Another example of a biological practice-based metaphysics can be found in the work of Thomas Reydon (2008), who has studied the ontological implications of biological practices in different disciplines, such as evolutionary and systematic biology. Particularly, he focuses on the use of the term 'species' in the scientific practice of these disciplines and concludes that, in evolutionary biology, the term species refers to an ontological 3D entity (evolveron), whereas in systematic biology it refers to an ontological 4D one (phylon).

I consider this form of interaction between metaphysics and biology to be different from the one found in Metaphysics *for* Biology. As noted above, in Metaphysics *for* Biology philosophers of biology address conceptual issues in biology by appealing to metaphysical theories and concepts that help to clarify the ontological status of the entities biological concepts refer to. In this case, therefore, the interaction goes from metaphysics to biology (see the red arrow in Figure 1), insofar as metaphysical concepts and theories are directly used to clarify a conceptual issue in biology. Conversely, in Metaphysics *in* Biology, metaphysics is not *directly* appealed when exploring the ontological commitments that might follow from biological theories, phenomena, and practices. Rather, these ontological commitments are obtained by directly considering biological theories, phenomena, and practices. In this case, the direction of interaction goes from biology to metaphysics (see the red arrow in Figure 2).

In some sense, this form of metaphysics seems to be the one that is usually recognized by the advocates of naturalized metaphysics as the only one that properly interacts with science (§ 1.2). This is, in fact, the kind of metaphysics that some philosophers of biology explicitly recognize as the one they do (Guay and Pradeu 2017; Dupré and Nicholson 2018). However, I do not consider naturalized metaphysics as the one that performs a role in this interaction.

As we saw in section 1.2, naturalized metaphysics is said to offer an ontological worldview by being constrained by our current sciences. Naturalized metaphysics is given insofar as the methodological, epistemological, and ontological constraints of science are followed (Soto 2015). However, naturalized metaphysics has been questioned for not being able to offer an accurate ontological worldview.

The problem on ontological plurality can also be seen within the framework of the philosophy of biology. Let's consider the current debate concerning whether biological phenomena imply an ontology of processes or one of substances. As previously explained, although some authors currently advocate that an ontology of processes is the accurate one when accounting for biological phenomena (Dupré 2012, 2015), other authors claim that an ontology of substances can also be derived from biological phenomena (Austin 2016b). By merely following the constraints of the biological phenomena and theories, it is not possible to decide whether an ontology of processes or one of substances is the correct one. In order to argue for a particular worldview, the participants in this debate need to appeal to arguments beyond biology, namely that offered by metaphysics. Furthermore, when philosophers of biology consider whether the ontology implied by the biological phenomena under study is one of substances or one of processes, they already have certain knowledge concerning what a substance and a process is. This idea is well illustrated in the following paragraph by Dupré and Nicholson:

> What is the difference between a thing, or a substance, and a process? [...] A starting point [...], is the following. Processes are extended in time: they have temporal parts. Whether things have temporal parts is a debated issue. [...] We shall not say any more about this debate here, except to note that four-dimensionalism is already halfway towards a full-blown process ontology. Both populate the world with temporally extended entities with diverse temporal parts (Dupré and Nicholson 2018, 11-12).

The metaphysical characterization of processes as temporally extended and having temporal parts does not follow from the mere study of biological phenomena, theories, and practices. Rather, it has been previously developed by metaphysicians addressing questions that concern, mainly, persistence. This kind of questions that are part of metaphysics might not be (and usually they are not) addressed in dialogue with any particular science. Thus, when philosophers of biology explore the metaphysical implications that follow from biological theories, phenomena, and practices, they do not merely follow the constraints of current biology, but they already have in mind metaphysical theories and concepts that are independent of science and that guide their work (see the dotted dark blue arrow in Figure 2).

In claiming that philosophers of biology do consider metaphysical concepts and theories when exploring the metaphysical implications and consequences that follow from biological theories, phenomena, and practices, I am not saying that the kind of metaphysics one finds in this interaction is the a priori one. In this form of interaction, philosophers of biology proceed by directly considering and paying attention to biological sciences. They do not a priori propose possible ontological worldviews in order to see, a posteriori, whether biological theories, practices or phenomena match some of these theories.

Since the philosopher of biology is appealing to the metaphysical toolbox in order to obtain the resources that allow her to explore the metaphysical implications that follow from biological theories, phenomena, and practices, I argue that the kind of metaphysics that plays a role in this case is the one I have termed applied metaphysics.

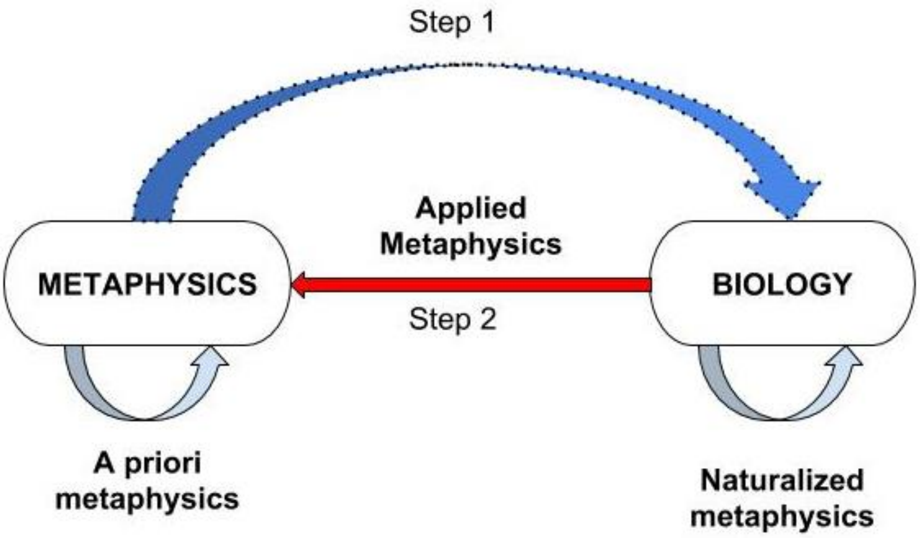

Figure 2. Metaphysics *in* Biology. The dotted dark blue arrow illustrates the first, previous step for the interaction between metaphysics and biology in this case. It refers to the already given metaphysical baggage and knowledge the philosopher of biology has when approaching biological theories, phenomena, and practices. Step two is illustrated by the red arrow, and it refers to the step in which the interaction between metaphysics and biology is properly given. In particular, it refers to the philosopher of biology's clarification of the ontological implications and worldview that follow from biological theories, phenomena, and practices. In step 2, the kind of metaphysics the philosopher of biology is performing is the one I call 'applied metaphysics'. It is not the a priori nor the naturalized one, represented by the light blue arrows. It is not a priori since philosophers of biology do not proceed by establishing a possible ontological worldview or implication that might follow from biological theories, phenomena, and practices without considering them. It is not naturalized either since the ontological implications and worldview that might follow from biological theories, phenomena and practices are not established by completely following the epistemological, methodological, and ontological criteria of biology, but by considering a previous metaphysical knowledge (illustrated by step 1) that is necessary for the philosopher of biology to properly proceed in this case.

4. The complexity of the interaction between metaphysics and biology: cross-fertilization and fuzziness

In the previous section, I have distinguished Metaphysics *for* Biology and Metaphysics *in* Biology as two different ways in which metaphysics and biology interact. Nonetheless, this distinction does not mean that these two kinds of interaction are neatly demarcated. Instead, in this section, I discuss how, in some cases, cross-fertilization and fuzziness characterize the interaction between metaphysics and biology.

4.1. Cross-fertilization

In some cases of interaction between metaphysics and biology, it is possible to find cross-fertilization between both disciplinary approaches (Reydon 2008). This occurs when the results obtained by philosophers in metaphysically addressing conceptual problems in biology also affect the metaphysical theories or concepts initially used.

This idea of cross-fertilization can be illustrated by means of an example. Let's consider Reydon's approach to the species problem. As previously seen (§ 3.2), Reydon argues that evolutionary and systematic biologists are referring to different ontological entities when using the term 'species'. Thus, evolutionary biologists refer to 3D ontological entities (evolverons) whereas systematic biologists refer to 4D ontological ones (phylons). Generally, in the contemporary metaphysical literature concerning persistence, 3D entities are equated to endurantism -in general terms, the theory that defends that entities persist by being wholly present at each time at which they exist; whereas 4D entities are equated to perdurantism -in general terms, the theory according to which entities are composed of different temporal parts, such that they persist by means of each temporal part being manifested at a different time at which the entity exists.[7] Since Reydon

---

[7] The question concerning the ontological status of the entities that exist in the world is different from the question about how these persist. However, endurantism and perdurantism tend to be associated with three-dimensionalism and four-dimensionalism, respectively (Miller 2005). As a result of this equation, in contemporary metaphysics, the

defends evolverons to be 3D entities and phylons to be 4D ones, he concludes that evolverons persist by endurance whereas phylons persist by perdurance (Reydon 2008).

This conclusion in philosophy of biology regarding the ontology of species might also affect the metaphysical theories and concepts initially considered by the philosopher, namely that of 3D (endurantism) and 4D (perdurantism). Reydon is aware of the cross-fertilization given in this case. In this sense, he states that to conceive of evolverons and phylons as persisting by endurance and perdurance, respectively, allows us to shed light on the metaphysical debate concerning the equivalence between both theories of persistence.

According to some metaphysicians, both theories of persistence endurantism (3D) and perdurantism (4D) are not really different theories but they are equivalent (see McCall and Lowe 2003, 2006; Miller 2005, 2010). On Reydon's view, however, the ontological consideration of species as 3D and 4D entities illustrates that such equivalence does not really hold (Reydon 2008; see Triviño and Cerezo 2015 for a different conclusion). As Reydon claims:

> For the species debate, the conclusion must be that the position that species three-dimensionalism and species four-dimensionalism constitute metaphysically equivalent theories is not a feasible choice. Note that the above consideration also shed some doubt on the general 3D/4D equivalence claim. The 3D/4D equivalence claim does not hold in the case of species, so it does not hold in general (Reydon 2008, 177).

---

debate between endurance and perdurance is translated into the debate between three-dimensionalism and four-dimensionalism.

Besides Reydon's case, other examples in the philosophy of biology literature illustrate the cross-fertilization that might result from the interaction between metaphysics and biology. One example refers to the case in Metaphysics *in* Biology concerning the ontological characterization of species as individuals that follows from Darwin's theory of evolution by natural selection. According to some authors, the ontological characterization of species as individuals has also affected the metaphysical characterization of individuality by showing that individuals do not need to be spatio-temporally cohesive entities (see Ghiselin 1997, ch. 4). As I will illustrate in the course of my thesis, the metaphysical approach to the problem of fitness is also an example of cross-fertilization between metaphysics and biology. In particular, as I will argue in Chapter 2, the application of the dispositional theory of causation to characterize the ontological status of fitness illustrates that this metaphysical theory can be improved by incorporating different relevant aspects of biological organisms.

4.2. Fuzziness

As previously stated (§ 3), the two kinds of interactions between metaphysics and biology proposed are different. Metaphysics *for* Biology directly recurs to metaphysical theories and concepts for clarifying the ontological reality biological concepts refer to in order to shed light on a conceptual problem in biology, whereas Metaphysics *in* Biology addresses this kind of biological problems by exploring the metaphysical implications and consequences that might follow from biological theories, phenomena, and practices. In the first case, metaphysical theories and concepts are directly considered when addressing the theoretical and conceptual problems at stake in biology. In the second case, notwithstanding, the philosopher of biology considers metaphysical concepts and theories in a more indirect way to be able to explore the metaphysical implications and consequences of particular biological theories, phenomena, and practices. The fuzzy character of the interactions between metaphysics and biology relies, precisely, on the difficulty to properly claim whether the appeal to metaphysics is direct or indirect in each particular case.

Coming back to the case of fitness, as previously stated, in order to clarify or shed some light on the problem of fitness in classic evolutionary biology, some philosophers of biology have appealed to different metaphysical concepts and theories such as that of dispositions or supervenience. In this regard, someone might claim that Metaphysics *for* Biology is the form of interaction given in this case. However, the problem of fitness could also be approached by paying attention to the ontological implications that follow from the biological work on fitness. In this sense, according to Darwin's theory, fitness is an ability of organisms to survive and reproduce. The notion of ability is considered to be, at least in some cases, synonymous of dispositions (Mumford and Anjum 2011; but see Maier 2014). In this sense, the metaphysical characterization of fitness as a disposition could be understood as something that follows from a biological theory, i.e. the theory of natural selection. Therefore, the kind of interaction given here would not be Metaphysics *for* Biology but rather Metaphysics *in* Biology.

As I consider, which form of interaction is the one given in each particular case is a question of emphasis. In cases of Metaphysics *for* Biology, the emphasis is on metaphysical theories and concepts, and how they could be useful for clarifying a conceptual and/or theoretical problem at stake in biology. Conversely, in cases of Metaphysics *in* Biology, the emphasis is on biological theories, phenomena, and practices. Metaphysical theories and concepts are not, at least in some cases, explicitly discussed, but they seem to be implied in the way the philosopher of biology proceeds. Moreover, there are also cases in which both kinds of interactions might be given. That is, cases in which it is possible to explore the ontological implications that follow from biological theories, phenomena, and practices, and then go deeper in clarifying those implications by directly appealing to metaphysical concepts and theories.

5. Towards a metaphysics of biology

As previously stated (§ 2), it is currently acknowledged (Guay and Pradeu 2017; Waters 2017; Dupré and Nicholson 2018) that philosophers of biology do not only appeal to methodological and epistemological frameworks to address conceptual problems in biology, but that a metaphysical approach is also endorsed in order to clarify this kind of problems (e.g. Boogerd et al. 2005; Huneman 2010; Dupré 2012, 2015; Austin 2016b, 2017; Waters 2017; Nicholson and Dupré 2018). This increasing tendency within philosophy of biology to recur to metaphysics has led some philosophers to talk about Metaphysics of Biology (henceforth MOB) (Guay and Pradeu 2017). In agreeing with this characterization, I believe my distinction between Metaphysics *for* Biology and Metaphysics *in* Biology, insofar as referring to two different ways MOB can be given, contribute to a more complex understanding of MOB.

In this regard, someone can claim that MOB is not a different kind of approach within philosophy of biology, but it refers to the study of metaphysical questions that appear in relation to a particular scientific discipline, i.e. biology. In this sense, MOB would refer to a particular case of

MOS, namely the one that Papineau calls 'particular metaphysical questions' (Papineau 1996) (§ 1). This characterization could be possible, although with a clarification. Recall that Papineau's characterization of MOS is the initial one in which MOS refers to the study of both metaphysical concepts that are common to all sciences and particular metaphysical questions that arise within a given science. However, this is not the way MOS is generally considered in current debates since it refers to the study of the *interaction* between metaphysics and science. In this sense, MOB could be considered as a particular case of MOS, but insofar as it refers to a particular case in which metaphysics and science interact. Thus, I think it is more appropriate to consider MOB as a new tendency within philosophy of biology, one in which philosophers of biology approach conceptual problems in biology from a metaphysical perspective besides the epistemological and methodological ones.

Insofar as MOB is a particular case in which metaphysics and science interact, the study of how this interaction takes place can also shed light on the more general debate in MOS. In this regard, some of the conclusions obtained in MOB can be extended to MOS. Firstly, there are different ways metaphysics and science can interact. Philosophers of science might metaphysically approach sciences in order to clarify the metaphysical status of the reality scientific concepts refer to. And they can also approach sciences in order to study the metaphysical implications that follow from scientific theories, practices, and phenomena. In both cases, the results allow for a clarification of an initial conceptual problem at stake in the particular scientific discipline philosophers consider. However, these approaches, notwithstanding the cases of fuzziness, are qualitatively different. The way the MOS debate is presented in the literature seems to indicate that the only kind of interaction that can be given is the second one, that is, that referring to the study of the metaphysical implications of scientific theories, phenomena, and practices.

Secondly, the interaction between metaphysics and science seems to be more complex than usually presented in the dichotomy *metaphysics prevalence view-science prevalence view*. From the particular case of MOB and the different interactions that can be recognized in it, it seems that both metaphysics and science are equally interacting and contributing in presenting an ontological worldview. Neither metaphysics guides science, nor science restricts or guides metaphysics. Furthermore, when a particular interaction between metaphysics and science takes place, cases of cross-fertilization (i.e. cases in which the results obtained when metaphysically approaching a particular science also serve to modify or improve the metaphysical theories initially used), might be given. In this sense, it is not only metaphysics that plays a role in clarifying a conceptual or theoretical problem in science, but science can also play a role in improving or modifying the metaphysical concepts and theories initially used.

Finally, it follows from my analysis of MOB that the form of metaphysics that plays a role in the interaction between metaphysics and biology, and therefore between metaphysics and science, is not the a priori nor the naturalized one. In this case, philosophers of science do not a priori establish possible ontological worldviews in order to check which one matches science better. Nor do they merely follow epistemological, methodological, and ontological constraints of science in order to obtain an ontological worldview. Instead, philosophers of science seem to appeal to the metaphysical toolbox in order to find the resources that properly allow them to perform their task, namely to shed light on some ontological features of the world.

In some cases, philosophers appeal to metaphysics in an indirect form, such that they have the metaphysical resources needed to obtain the ontological implications that follow from scientific results without making them necessarily explicit. This could be, for instance, Ladyman and Ross' case concerning ontic structural realism since, as already stated by French and McKenzie (2015), this ontological view does not directly follow from quantum mechanics, but it presupposes metaphysical concepts and commitments such as that regarding the characterization of the fundamental, the notion of structure, or the relation between the fundamental and the non-fundamental. Yet, in other cases, philosophers of science might use metaphysics in a more direct and explicit way when clarifying the ontological reality scientific concepts refer to. This situation might be illustrated by the use of the metaphysical concept of disposition (and the different metaphysical theories about them) to characterize the ontological reality quantum mechanics refer to (Dorato and Esfeld 2010).

The idea of a metaphysical toolbox that philosophers of science use when approaching science from a metaphysical perspective has been proposed by French and McKenzie (2012, 2015). They claim that the form of metaphysics that is part of the toolbox is that performed by metaphysicians without taking into account the scientific knowledge or scientific results, that is, the metaphysical toolbox is composed of those theories and concepts that are considered to be part of analytic metaphysics.[8] It is the philosopher of

---

[8] I maintain here French and McKenzie's concept of analytic metaphysics in order to distinguish it from the characterization of a priori metaphysics that has been used in this chapter and that is the common one when characterizing the MOS debate. In this sense, a priori metaphysics shares with analytic metaphysics its scientific-independent way of

science who decides whether a particular metaphysical theory or concept is useful or not for addressing a theoretical or conceptual issue in science.

Analytic metaphysics, therefore, does not need to be concerned with science. At least, it is not as long as it does not try to describe the ontological reality adequately (Le Bihan and Barton 2018). In order to describe the ontological reality, both metaphysics and science are necessary. According to French and McKenzie, the fact that the philosopher of science might use some metaphysical concepts and theories to approach conceptual issues in science and to obtain an ontological worldview justifies the heuristic value of analytic metaphysics (French and McKenzie 2012, 2015). I agree with French and McKenzie in this regard. Furthermore, I consider that in the task of appealing to metaphysical concepts and theories to obtain a metaphysical worldview or to approach science, the philosopher of science is doing applied metaphysics, i.e. one that is not a priori (since it takes into account scientific theories and results when offering an ontological worldview), nor naturalized (since metaphysical concepts and theories are, at least, presupposed when exploring the ontological consequences that follow from scientific results).

French and McKenzie have claimed that the heuristic value of analytic metaphysics might be questioned on the basis that there could be metaphysical concepts and theories that might be never used by philosophers of science. In this regard, I agree with Baptiste Le Bihan and Adrien Barton (2018) when claiming that science is continuously changing and thus, metaphysical concepts and theories that are not currently useful for philosophers of science could be so in the future. In this sense, they claim, the mere fact that a metaphysical concept or theory could be useful in the future

---

proceeding. However, it is different insofar as it tries to offer an ontological characterization of the world whereas analytic metaphysics does not try to do so.

justifies the current metaphysical work on it. In other words, analytic metaphysics has heuristic value insofar as metaphysical concepts and theories within it might be useful either at present or future times (Le Bihan and Barton 2018).

6. Conclusion

In this chapter, I have presented the philosophical framework of my thesis, namely that about the interaction between metaphysics and biology. In particular, I have characterized this interaction by taking into account the results of both my study of cases in the literature wherein an interaction between metaphysics and biology takes place, as well as my own work on conceptual and theoretical problems in biology that I have addressed from a metaphysical perspective. Based on these case studies, I have proposed two different ways metaphysics and biology seem to interact: Metaphysics *for* Biology and Metaphysics *in* Biology.

The characterization of these two kinds of interactions between metaphysics and biology is useful insofar as it allows us to shed some light on the clarification of an increasing tendency in philosophy of biology to address conceptual problems in biology from a metaphysical perspective, a tendency that I have agreed to term 'Metaphysics of Biology' (MOB) (Guay and Pradeu 2017). Furthermore, exploring the interaction between metaphysics and biology is also useful in approaching the general debate in MOS concerning the interaction between metaphysics and science. This is so since MOB represents a particular interaction between metaphysics and science. In this regard, three main outcomes have been considered and applied to the debate in MOS from the study of the interactions between metaphysics and biology.

Firstly, there are different ways metaphysics and science interact. Secondly, the relation between metaphysics and science seems to be more complex than usually characterized. It is not that metaphysics either guides or is a condition for the possibility of science, nor it is science that constraints metaphysics. Rather, both metaphysics and science seem to be equally interacting when philosophers try to obtain an ontological worldview from the sciences. And thirdly, the kind of metaphysics that seems to be playing a role in the interaction is not the a priori nor the naturalized one, but one I have labeled 'applied metaphysics'. In this regard, neither science nor metaphysics by themselves are able to offer an ontological worldview. Rather, both disciplines need to work together in performing this task.

In the rest of my thesis, I will present two conceptual problems in philosophy of biology that I have addressed from a metaphysical perspective, namely the problem of fitness and that concerning the ontological status of holobionts. As previously stated, both problems instantiate the kinds of interactions between metaphysics and biology that I have distinguished in this chapter.

Chapter 2:
Fitness as a Causal Disposition

> *"I look at all the species of the same genus as having as certainly descended from the same progenitor, as have the two sexes of any one of the species. Consequently, whatever part of the structure of the common progenitor, or of its early descendants, became variable; variations of this part would, it is highly probable, be taken advantage of by natural and sexual selection, in order to fit the several species to their several places in the economy of nature, and likewise to fit the two sexes of the same species to each other, or to fit the males and females to different habits of life, or the males to struggle with other males for the possession of the females."*
> (Darwin 1859)

Fitness is a key concept in evolutionary biology. It is used to explain evolutionary change: natural selection operates whenever individuals in a population differ in their fitness. In Darwin's original formulation, what we call 'fitness' referred to the relative *abilities* of individual organisms in a population to survive and reproduce. Currently, notwithstanding, biologists define fitness in terms of the actual number of offspring left by an individual or type.

This characterization of fitness is known as the actualist account of fitness, and it has been challenged for involving circular explanations and therefore, lacking explanatory power (Brandon 1978; Rosenberg 1985; Millstein 2016). Besides this problem, the actualist account has also been charged for not being able to distinguish evolution by natural selection from evolution by drift (Brandon 1978, 1990, 2005; Beatty 1984; Millstein 2002, 2005), and, in this sense, this definition of fitness does not coincide with the way biologists really use the notion (Scriven 1959; Mills and Beatty 1979).

In order to avoid the problems of the actualist account, philosophers of biology have proposed alternative characterizations of fitness. Some of these approaches to fitness have appealed to metaphysical concepts allowing to distinguish fitness from actual rates of reproduction. In the late 1970s, Alexander Rosenberg offered a supervenient approach in which fitness is considered to be a supervenient property of organisms whereas the actual number of offspring is only a way of measuring the level of fitness an organism has (Rosenberg 1978). Roughly at the same time, Robert Brandon (1978), as well as Susan Mills and John Beatty (1979), offered a propensity interpretation of fitness that conceptualizes fitness as an ability of organisms to survive and reproduce in a particular environment and population. This ability is expressed in probabilistic terms by appealing to the number of offspring an organism is expected to leave (expected fitness). This probability, notwithstanding, might not match the number of offspring the organism actually leaves (realized fitness). In this sense, the fitter organism is the one that has the best chance of outsurviving and outreproducing the less fit. Yet, it may not do so (Brandon and Beatty 1984; Beatty and Finsen 1989).

Rosenberg's supervenient account and Mills and Beatty's propensity interpretation of fitness are metaphysical approaches to fitness. They try to clarify the ontological status of fitness as a property of organisms. However, this first ontological characterization of fitness was very loose. In this sense, although Rosenberg and Mills and Beatty appeal to metaphysical concepts, they do not properly characterize the ontological status of fitness that follows from the metaphysical characterization they propose. Furthermore, as I will argue along this chapter, in both approaches, the mathematical form of measuring fitness seems to prevail over its ontological characterization. In this regard, philosophers of biology have highlighted the necessity to properly distinguish how fitness *ontologically* is from how it is measured (Brandon 1990; Millstein 2016).

Currently, some authors have worked on improving the ontological characterization of fitness, wherein my work in this chapter is also situated (Matthen and Ariew 2002; Bouchard and Rosenberg 2004; Abrams 2006, 2007). In particular, I will appeal to the metaphysical framework of the dispositional theory of causation recently proposed by Stephen Mumford and Rani Lill Anjum (2011) which, as I will argue, offers a variety of conceptual resources that allow for a clarification of the ontological status of fitness as a property of organisms. Furthermore, under this metaphysical framework, the problems of the actualist account of fitness can be solved and the distinctiveness that biologists usually attribute to fitness, that is, the fact that fitness differs from both the number of offspring an organism has and the traits that characterize it, are better explained.

The structure of the chapter is as follows. Firstly, I briefly introduce the problem of fitness in classic evolutionary biology (§ 1). Secondly, I focus on some metaphysical approaches to fitness proposed by philosophers of biology (§ 2). In particular, I will pay attention to the supervenient account (§ 2.1) and the propensity interpretation of fitness (§ 2.2), which is considered to be one of the most promising proposals allowing to characterize fitness while differentiating it from its measurement. After showing some of the deficiencies of the current propensity interpretation of fitness, I briefly introduce the metaphysical debate concerning the ontology of properties (§ 3). In Section 3.1, I focus on the characterization of categorical and dispositional properties in order to introduce the metaphysical framework of

Mumford and Anjum's dispositional theory of causation (§ 3.2). As I will argue (§ 4), a more accurate approach to the ontological status of fitness can be made within this framework. In this regard, I characterize fitness as a causal disposition that is complex-dynamical (§ 4.1) and context-sensitive (§ 4.2). Finally (§ 4.3), I illustrate how the notions of fitness, expected fitness, and realized fitness are distinguished in this account.

1. The problem of fitness in classic evolutionary biology

The notion of fitness is a key concept invoked in classic evolutionary biology to explain evolutionary change: natural selection operates whenever individuals in a population differ in their fitness. The original formulation of the concept can be found in Darwin, although he did not explicitly used the term 'fitness' but several similar terms such as 'fit', 'fitting' or 'fitted' (Millstein 2016). In Darwin's formulation, what we call 'fitness' referred to the *disposition* or *ability* of individual organisms to survive and reproduce. Thus, fit organisms are those more able to survive and reproduce in a particular environment and population. For Darwin, fitness is a relational term. It is in relation to a given environmental and populational context that organisms are said to be fitter or less fit than others. An individual organism, therefore, cannot be fit in an absolute sense (Millstein 2016). Since, in this formulation, fitness is understood as an ability or disposition of organisms, fitter organisms are those that have a better chance of surviving and reproducing. Yet, this characterization leaves open the possibility for the less fit organism to outsurvive and outreproduce the fittest one (Brandon 1978).

The original Darwinian characterization of fitness experienced a change during the early 20th century due to the mathematization of the evolutionary theory undertaken by population genetics. The emergence of theoretical population genetics is associated with the work of Ronald Fisher, *The Genetical Theory of Natural Selection* (1930), Sewall Wright, "Evolution in Mendelian Populations" (1931), and John Burdon Sanderson Haldane, *The Causes of Evolution* (1932). Population genetics allowed the mathematical reconciliation of Mendelism with Darwinism and became the conceptual core of the theory of evolution as articulated by the so-called Modern Synthesis of Evolutionary Biology (Huxley 1942).

Population genetics is addressed to understand, mainly, the genetic architecture and the origin of the so-called 'adaptive norm', that is, the idea that populations are in a relation of adaptedness (fitness) to its environment (Dobzhansky 1954). As Theodosius Dobzhansky has made explicit:

> An array of related genotypes consonant with the demands of the environment is the adaptive norm of a population [...]. The gene pool of every Mendelian population contains a greater or lesser variety of phenotypes, and every population meets a variety of environments. In contrast to the adaptive norm, every large population contains some genotypes and phenotypes which are not competent to survive and reproduce in at least some of the environments which the population has to face [...] The adaptive norm of a Mendelian population is a product of its evolutionary history (Dobzhansky 1954, 3).

Population genetics is a subfield of evolutionary biology. It conceives evolution as a change in the alleles frequencies of genes in a population over time. In this regard, evolution is mathematically represented by measuring changes in genetic variation of populations over time due to the influence of various causal factors (i.e. mutation, migration, natural selection, random drift) in those populations.[9] In particular, population genetics focuses on the so-called 'Mendelian populations', that is, breeding populations of sexually reproducing species (Dobzhansky 1954; Millstein and Skipper 2007).

Concerning the causes of evolution, the one that population geneticists more widely consider is natural selection, since it is natural selection that selects a particular combination of alleles (Gould 1993).[10] Fitter organisms, i.e. those with better genetic combinations, would have more offspring, and thus, populations can reach a temporary genetic stability corresponding with the adaptive norm of the population in the particular environment it is embedded in.

In population genetics, fitness is a mathematical parameter within the Hardy-Weinberg equation that refers to the number of offspring of an organism. It is measured in a comparative way: by paying attention to the number of offspring left by an organism within a population as compared to others (Millstein and Skipper 2007). The Hardy-Weinberg equation specifies

---

[9] There is a debate regarding whether natural selection is, in fact, a cause of evolution, and whether it can be distinguished from genetic drift. I will not pay attention to this issue since it goes beyond the purposes of this thesis. For further detail, see Beatty 1984; Hodge 1987; Millstein 2002; and Millstein and Skipper 2007.

[10] In this regard, it is important to notice that there are two forms of conceiving natural selection as being a cause of evolution (Millstein and Skipper 2007): 1. Natural selection is a cause of evolution itself, i.e., natural selection is a force that acts in population and produces changes in its genotypic variation (Sober 1993). And 2. Natural selection is a cause of evolution insofar as it is made up of causes. In this latter sense, fitness is considered to be a cause within natural selection. It is an organism fitness that causes it to have a survival and reproductive success in a particular environment and population. Both senses are tightly connected. In fact, whether one considers natural selection to be, by itself, a cause of evolution, depends on whether one considers fitness as a cause. In this regard, philosophers of biology are divided into causalists, i.e. those who argue for the causal power of fitness (Brandon 1978, 1990; Mills and Beatty 1979; Pence and Ramsey 2015; Ramsey 2015), and statisticalists, i.e. those claiming that fitness is a mere statistical summary of underlying causal processes (Sober 1993; Matthen and Ariew 2002; Walsh, Lewens and Ariew 2002; Walsh 2007, 2010; Saphiro and Sober 2007).

the conditions under which evolution does not occur, i.e. under which there is no change in the gene frequencies of a population from one generation to the next. In those cases in which evolution does not occur, the population is said to be in Hardy-Weinberg equilibrium (Millstein and Skipper 2007). By knowing the distribution of allele frequencies in the absence of evolutionary causes, population geneticists are able to calculate the effects of a given cause or causes acting in a population by measuring its influence in the distribution of those allele frequencies.

Due to its mathematical use, biologists started to equate fitness to the actual number of offspring left by an organism. In this regard, John Marion Thoday, a biologist, spoke of fitness in 1953 in the following way:

> Therefore the fitness of each generation is dependent to some measure on the mates of its descendants and we are forced to define fitness for survival in relation to that contemporary group of individuals which will some time in the future have common descendants (Thoday 1953, 97).

The so-called 'actualist account of fitness' (Millstein 2016), notwithstanding, has led to different problems that have been highlighted by both biologists and philosophers of biology. The main one is that this account makes evolutionary explanations based on fitness to be circular. This is the so-called *tautological problem* (see Millstein 2016, for a recent review). If we seek to explain why an organism *A* leaves more offspring than an organism *B*, the explanation "because *A* is fitter than *B*" is circular if "*A* is fitter than *B*" means "*A* leaves more offspring than *B*" (Brandon 1978; Rosenberg 1985). In other words, if the fittest organisms are the reproductively successful ones, then those reproductively successful organisms are reproductively successful (Burian 1983). The tautological problem illustrates that the actualist account of fitness lacks explanatory power due to its conceptual connection with rates of reproduction. As Mills and Beatty state it:

> The fact that references to fitness lead to explanatory circularity just shows that fitness has no explanatory role to play in evolutionary biology. In fact, Bethell (1976, p.75) makes this latter claim, and even maintains that biologists have abandoned references to fitness in their accounts of evolutionary phenomena. This is a scandalous claim (Mills and Beatty 1979, 267).

Besides the tautological problem, the actualist account has also been criticized for not capturing the real *use* that biologists make of the concept of fitness in their scientific practice. Biologists do not seem to use the concept of fitness in the same way as they define it. This is the so-called *mismatch problem* (Mills and Beatty 1979; Millstein 2016). The most celebrated example to illustrate this mismatch is the one developed by Scriven (1959). Let's consider two identical twins in a forest. One of them, is struck by lightning such that it dies and does not leave offspring. The other organism, notwithstanding, is spared and reproduces. In terms of the actualist account of fitness, the lucky twin is the fitter one, insofar as it has offspring whereas the other has not. Yet, this claim concerning fitness seems counterintuitive. In fact, most biologists would agree in considering that both organisms are equally fit given that both are assumed to be phenotypically and genetically identical.

The mismatch problem stems from a more general deficiency of the actualist account, namely it is unable to offer a criterion to distinguish evolution resulting from differences in fitness from evolution that occurs by drift, that is, changes in allele frequencies of a population that are due to chance (sampling error), and not to natural selection. I will refer to this problem as *the lack of criterion problem*. Accidents might occur in nature such that an organism might outreproduce another one without being the fittest. Yet, since fitness refers to the actual reproductive outcome of organisms, biologists would not be able to distinguish between this case of chance and a case of natural selection, that is, a case in which differences in the actual number of offspring are caused by some phenotypic difference in organisms. As Brandon has explicitly claimed:

> Natural selection is not just differential reproduction, but rather it is differential reproduction that is due to differential adaptedness, that is, due to the adaptive superiority of those who leave more offspring (Brandon 2010, 111).

In order to avoid the problems referred above concerning the actualist account, biologists have offered a wide amount of alternative definitions of fitness (Barker 2009). However, this conceptual clarification is not an easy task. As Dobzhansky has claimed, both adaptedness (fitness) and adaptation are among the fundamental biological concepts for which no precise definitions can yet be given (Dobzhansky 1954). More recently, the biologist Harold Allen Orr has explicitly claimed that the problems concerning fitness are more conceptual than empirical since it is often more difficult to think clearly about fitness than to perform experiments on it (Orr 2009).

In the late 1970s, philosophers of biology also started to develop alternative approaches to fitness which allow to distinguish it from actual rates of reproduction. Their aim concerning this task was not only to solve the problems of the actualist account of fitness but to offer an alternative definition of the concept that could be useful for biologists and that allowed the clarification of some of the claims in the evolutionary theory (Mills and Beatty 1979). The debate in philosophy of biology concerning fitness has also been integrated into biology since both biologists and philosophers are interested in the conceptual clarification of this notion. In fact, it is not rare to find biologists that use definitions of fitness proposed by philosophers of biology. Orr, insofar as assuming Brandon's characterization of fitness as an ability of organisms to survive and reproduce in their environment and population (Brandon 1978), is an example of this interaction (Orr 2009).

2. The problem of fitness in philosophy of biology

In this section, I will present some of the accounts of fitness offered by philosophers of biology. In particular, I will focus on Rosenberg's supervenient account of fitness (Rosenberg 1978) (§ 2.1), as well as in Brandon's (1978) and Mills and Beatty's propensity interpretation of fitness (1979) (§ 2.2), which is the more widely used and acknowledged due to its conceptual richness (Millstein 2016).

2.1. The supervenient account of fitness

In order to conceptually clarify the notion of fitness, Rosenberg developed a supervenient approach that makes possible to distinguish fitness from rates of reproduction (Rosenberg 1978). In his account, fitness (i.e. levels of fitness) refers to the real physical properties that characterize organisms, whereas rates of reproduction (i.e. the number of offspring an organism leaves) are a way of measuring the particular level of fitness an organism has. Thus, levels of fitness should not be equated to rates of reproduction: what a thing is should not be equated to how that same thing is measured.

In Rosenberg's view, the level of fitness of an organism is based on the anatomical, physiological, behavioral, and environmentally relative properties that characterize such organism (Rosenberg 1978, 372). To properly calculate the level of fitness of an organism, therefore, implies to consider the different properties that characterize it. However, Rosenberg argues, these properties are so varied and difficult to determine that biologists tend to appeal to rates of reproduction in order to calculate a given level of fitness (Rosenberg 1983). Rates of reproduction are a form of measuring the level of fitness of an organism since they are the *effect* of the organism having a particular level of fitness. In other words, the properties that characterize an organism *cause* it to leave a certain number of offspring. As Rosenberg claims:

> [...] the properties of organisms [...] govern an organism's number of reproductive opportunities, and rate of successful reproduction. They are, in short, the causes of its rate of reproduction. Naturally, the instantiation of these properties causally connected to organisms' rates of reproduction will themselves be explained by anatomical, physiological, behavioral and ecological theories (Rosenberg 1978, 373).

Appealing to rates of reproduction, therefore, is useful insofar as we do not always have access to all the properties upon which the level of fitness of an organism depends. Furthermore, rates of reproduction also allow us to calculate the level of fitness of two organisms that greatly differ in their properties. Rosenberg illustrates this situation by means of a bird and a squirrel that "can occupy roughly the same environment, have exactly the same prospective and retrospective reproduction rates, and yet differ greatly in their properties" (Rosenberg 1978, 371). Since these organisms have exactly the same reproduction rates, it is possible to claim that both have the same level of fitness despite differing in their properties. Yet, as previously noticed, although rates of reproduction refer to a way of measuring levels of fitness, the particular level of fitness of an organism should not be equated to its number of offspring (Rosenberg 1983, 460).

Since levels of fitness are based on the physical and environmentally relative properties of organisms, two organisms that are alike regarding these properties would also have the same level of fitness, although they might have different rates of reproduction (Rosenberg 1978, 374). This is the case of the two twins standing in a forest. Insofar as both organisms are identical, both have the same level of fitness in spite of the fact that one of them does not leave any offspring. Rosenberg conceptualizes this idea by considering levels of fitness to be *supervenient* on "those properties, dispositions, and abilities which organisms have in virtue of their anatomical, physiological character, and the interaction of this character with the organism's environment" (Rosenberg 1978, 372).

Supervenience is a term of art in philosophy that has its origins in the work of the British Moral Theorists (Depaul 1987; Kim 1990; Beckermann 1992). Since its origins in the moral field, supervenience refers to a relation of covariation between two sets of properties, i.e. the lower-level (subvenient) properties and the (higher-level) supervenient ones (Kim 1982, 1984, 1990). According to this kind of relation, there cannot be two entities that are alike regarding their subvenient properties but different in some supervenient property, in other words, an object cannot vary in some supervenient property without a variation in some subvenient property. By claiming that the level of fitness supervenes on the properties of organisms, Rosenberg is able to explain both that organisms with the same properties do have the same level of fitness, and that the same level of fitness might be given in organisms with different properties, such as a bird and a squirrel. The only requirement, in this case, is that a change cannot be produced regarding the level of fitness of any of these organisms without also being a change in the lower-level properties that characterize them.

Due to the supervenient character of levels of fitness, Rosenberg argues that fitness is a *supervenient concept*. In this regard, fitness is a useful concept insofar as it plays an explanatory role, but it does not refer to something different from the properties that characterize organisms. Furthermore, and since the traits of organisms that might contribute to (i.e. cause) their survival and reproduction are endlessly varied and context-dependent, it is not possible to identify fitness with a particular physical property of organisms. As Elliott Sober makes explicit in a reply to Rosenberg, fitness is not a physical property of organisms, although each organism's fitness consists in the constellation of its physical properties (Sober 1984, 376). According to Rosenberg, this peculiarity of fitness illustrates that it is a *primitive* and *undefinable* term within the theory of evolution (Rosenberg 1983; see also Williams 1970). It is used for simplicity, in order to avoid the enumeration of all the traits of organisms that are connected to their survival and reproductive success.

2.2. The propensity interpretation of fitness

Another alternative to the actualist definition of fitness was independently and simultaneously proposed by Brandon and Mills and Beatty. According to the so-called propensity interpretation of fitness (henceforth PIF) (Brandon 1978; Mills and Beatty 1979), fitness (or 'adaptedness' in Brandon's terminology) does not refer to the actual survival of an organism or the number of offspring an organism actually leaves, but to the propensity, ability, or disposition[11] the organism has to survive and reproduce in a

---

[11] Philosophical terms such as that of disposition, ability, propensity, capacity, etc., might have different meanings in different philosophical and metaphysical fields, such as in

particular environment and population. This ability of organisms is expressed by means of the number of offspring an organism is expected to have in a particular environment and population, what is referred to as the expected fitness. The expected fitness can differ from the so-called realized fitness, which refer to the actual reproductive success of organisms.

It is important to note that this account is compatible with the supervenient one. As Sober (1984, 1993) claims, although fitness can be considered as a disposition, it is one of a special kind since it can be multiply-realized insofar as it supervenes on different physical and chemical properties (see also van der Steen 1994). However, the conceptual resources PIF introduces (such as the distinction between expected and realized fitness), as well as the characterization of fitness as a disposition, allow us to make more sense of the way biologists use the concept and to solve the problems of the actualist account in a more accurate way.

Regarding the theory of evolution, Mills and Beatty reject to consider fitness as a primitive and undefinable term (as per Rosenberg). As they highlight, evolutionary biologists continuously appeal to the concept of fitness in order to offer explanations of evolutionary phenomena. And, in doing so, they try to refine and extend the concept. However, Mills and Beatty claim, biologists have difficulties in defining fitness since they confuse "the *post facto* survival and reproductive success of an organism, with the *ability* of an organism to survive and reproduce" (Mills and Beatty 1979, 270). But fitness, they argue, refers precisely to the ability of an organism to survive and reproduce in a particular environment and population (Mills and Beatty 1979).

---

probability theory. However, in the debates concerning the definition of fitness, they are used interchangeably and considered as synonymous.

By considering the dispositional character of fitness, Mills and Beatty claim to recover the main properties present in the original Darwinian conception of fitness, namely it is an ability of organisms; it is relative to an environment and population; and it refers to both survival and reproduction (Millstein 2016). Furthermore, the characterization of fitness as a disposition also accounts for its explanatory power. To illustrate this idea, Mills and Beatty compare the disposition of fitness with the disposition of solubility. Fitness, insofar as it is a disposition, explains why a particular organism survives and reproduces in its environment and population in the same way as solubility explains that a certain substance, such as a cube of salt, dissolved when introduced in water. The actual survival and reproduction of an organism, just like the dissolution of salt, refer to the manifestation associated with a certain disposition. Thus, it is the possession of the disposition (i.e. fitness) that explains the manifestation it produces (i.e. to survive and reproduce).

Concerning the ontological status of dispositions, Mills and Beatty are reductionists, that is, they are not ontologically committed to the existence of dispositional properties (I will say more on this in § 3). As they make explicit, dispositional properties of an entity can be reduced to its physical, i.e. categorical, properties, which are the ones causally relevant in producing an effect once certain triggering conditions are given. Within this reductionistic framework, therefore, the disposition of salt to be dissolved in water refers to its ionic crystalline character since it is this character that really causes the dissolution of salt when the triggering conditions are met, that is, when salt is introduced in water. In the same way, the disposition of fitness can be reduced to the traits that characterize organisms and that are the causes of their survival and reproduction when the triggering conditions are given. In this case, the triggering conditions refer to the particular environmental conditions wherein the organism is embedded (Mills and Beatty 1979, 271).

In the same line as Mills and Beatty, Brandon does also consider fitness as a disposition that can be reduced to the traits of organisms. In his view, certain traits will render an organism *better adapted* to its environment than a conspecific with certain other traits (Brandon 1978, 1990). By considering fitness as an ability that is grounded on the physical traits of organisms, the proponents of PIF also recognize the causal power of fitness besides its explanatory role. Thus, "[a]ccording to the propensity interpretation, *the connection* between the ability and the actual manifestation of the ability *is a causal connection*" (Brandon and Beatty 1984, 344. Emphasis added).

In metaphysics, one of the core features of dispositions is that they do not always have to reach their manifestation despite the triggering conditions being met. Some dispositions might be *masked* (Johnston 1992) so that they are not triggered regardless of the presence of triggering conditions, whereas others might be *prevented* (Bird 1998a) from manifesting despite having been triggered. An example of the first case, masked dispositions, is a cube of salt that is covered by plastic. In this case, solubility will not manifest despite the introduction of the cube of salt in water. The second case, prevention, is usually illustrated by appealing to antidotes (Bird 1998a). In those cases wherein someone swallows a poison, it is possible to stop the manifestation of the poison (i.e. to kill) by taking some kind of antidote. Mills and Beatty are aware of these metaphysical considerations. Yet, regarding fitness, they make clear that possible factors that might interfere with its manifestation are not taken into account:

> Likewise, when we explain an organism's (or type's) offspring contribution by referring to its degree of fitness, we assume, for

> instance, that environmental catastrophes (e.g. atomic holocausts, forest fires, etc.) and human intervention have not interfered with the manifestation offspring contribution dispositions. In general, we want to rule out the occurrence of any environmental conditions which separate successful from unsuccessful reproducers without regard to physical differences between them (Mills and Beatty 1979, 272).

Due to its dispositional nature, Mills and Beatty hold that fitness is a property of organisms: dispositions are not properties of classes or abstract entities, but of individual entities. However, they recognize that it is also possible to talk about the fitness of types, as referring to phenotypes and/or genotypes. In these cases, fitness should not be conceptualized as a propensity, but rather as a derivative of individual fitness propensities (Mills and Beatty 1979, 272).[12] It is individual fitness that allows to make comparisons among different individuals and to group them into types:

> [T]he evolutionary biologist is interested in explaining proportions of types in populations, and from this point of view, an organism which leaves one offspring is much more similar to an organism which leaves no offspring than it is to an organism which leaves 100 offspring (Mills and Beatty 1979, 272-273).

Concerning individual fitness, Mills and Beatty remark that, although fitness is a propensity of organisms, this propensity does not refer to whether the organism would reproduce or not in a given environment, but to the *number* of offspring the organism *might have* due to its propensity. In this sense, the propensity of organisms is a quantitative propensity. What is relevant for natural selection is not whether an organism leaves offspring or not, but how many offspring it contributes to the next generation.

---

[12] Since my main concern refers to the way *individual fitness* can be ontologically characterized, and since in Mills and Beatty account type fitness is derivative from individual fitness, I will only briefly present type fitness in this section.

According to Mills and Beatty, the quantitative character of individual fitness is non-deterministic, insofar as there is not "a unique number of offspring an organism is determined to leave". Rather, there are different propensities for the organism "to leave different numbers of offspring" (Mills and Beatty 1979, 273).

Based on these considerations, Mills and Beatty argue that the fitness of organisms (what they call *fitness*$_1$) can be measured in terms of the probabilistic distributions associated with their reproductive propensities, by assigning values to each propensity and calculating the weighted sum of these values. This mathematical result (that is, the average number of the different offsprings numbers associated with an organism's propensities to reproduce) will refer to the 'expected fitness' of organisms (Mills and Beatty 1979, 274). Expected fitness, therefore, is defined as the number of offspring that is expected for an organism to leave. By taking into account this quantitative character of fitness, Mills and Beatty define organismic fitness (*fitness*$_1$) as follows:

> The *fitness*$_1$ of an organism *x* in environment *E* equals *n* $=_{df}$ *n* is the expected number of descendants which *x* will leave in *E* (Mills and Beatty 1979, 275).

The quantification of the fitness of an organism, therefore, is given by means of its expected number of offspring, which is an indication of the organism's ability to survive and reproduce in a particular environment and population. The expected fitness of an organism, notwithstanding, might not coincide with the actual number of offspring it leaves (realized fitness). The expected number of offspring is calculated by taking into account the phenotypic traits that characterize the organism (determined by means of optimality or biomechanical models (Millstein 2016)) as well as certain initial environmental conditions. Yet, as previously stated, this calculus is made in *ideal conditions*, without considering other factors that might affect the manifestation of the fitness disposition, such as an environmental catastrophe. This kind of factors, notwithstanding, can occur, and they would explain that the organism with a higher expected fitness is not always the one that survives and reproduces the most.

On the basis of organismic fitness or *fitness*$_1$ Mills and Beatty formulate type fitness or *fitness*$_2$. In this case, fitness "reflects the contribution of a particular gene or trait to the expected descendent contribution (i.e. *fitness*$_1$) of possessors of the gene or trait" (Mills and Beatty 1979, 276). However, as Mills and Beatty note, it is not easy to calculate the contribution of a type of gene or trait insofar as organisms possessing the particular gene or trait will vary in relation to other genes and traits they possess and that do also affect their fitness. By taking this difficulty into consideration, Mills and Beatty propose to define type fitness as the average expected number of offspring of the members of the types (genotype or phenotype) under consideration, that is, the average *fitness*$_1$.

By distinguishing between fitness and actual reproductive success, PIF has important advantages over the actualist account of fitness, such as its capacity to overcome the tautological, the mismatch, and the lack of criterion problems. Concerning *the tautological problem*, when we seek to explain why organism *A* leaves a higher number of offspring than organism *B*, the explanation "because *A* is fitter than *B*" means that "*A* has a greater propensity than *B* to survive and reproduce in the given environment". This explanation is an accurate one insofar as it refers to the idea that the phenotypic properties of *A* in its environment cause it to tend to have greater reproductive success than *B*. As stated above, fitness is an organism's ability that is measured in terms of its expected fitness. The expected fitness is, thus, only an indication of the organism's ability to survive and reproduce: a fitter or better adapted organism has the best chance of outsurviving and outreproducing the less fit, but it may not do so. In this sense, it is possible for the less fit organism to actually have more offspring than the fitter one despite their differences in their ability to reproduce (Beatty and Finsen 1989, 17-18; Brandon and Beatty 1984). *The mismatch problem*, illustrated by the case of the two twins standing together in a forest, is also easily solved in this approach. In this case, biologists could claim that both twins are equally fit, insofar as their properties confer them the same disposition to survive and reproduce, that is, both have the same expected fitness. Yet, they differ in their realized fitness since the disposition is only realized in one case.

Finally, PIF is able to account for *the lack of criterion problem* as well. As previously stated, expected fitness is predicted by taking into account the organisms' phenotypic features. Environmental factors and conditions that might interfere with the manifestation of fitness are not considered. In this regard, Mills and Beatty claim that cases of natural selection are given in those situations wherein the expected and the actual number of offspring do actually coincide. This is so since expected fitness is calculated by means of the phenotypic properties of organisms, which are causally connected to their number of offspring. Conversely, in those cases in which expected and realized fitness do not coincide there has been evolution by chance, not by natural selection. In these cases, the number of offspring a particular organism has is not due to its phenotypic properties but to some environmental factor that interfered with the organism's fitness disposition. In PIF, therefore, cases of evolution resulting from differences in fitness (i.e. a discriminate sampling process in which physical differences between organisms of one generation are relevant to differences in their offspring contributions), and those resulting from drift (i.e. an indiscriminate sampling process in which physical differences between organisms of one generation might be irrelevant to differences in reproductive success) can be properly distinguished (Burian 1983; Beatty 1984, 189; Brandon 1990, 2005; Millstein 2002, 2005, 2016).

Despite all the virtues of PIF referred above, there are some aspects of this approach that are considered to need more clarification. One of these issues concerns the way probability is understood in this account and how the probabilistic quantification of fitness affects the consideration of evolution by natural selection as a deterministic or as an indeterministic process (Rosenberg 1985; Rosenberg and Williams 1986; Hodge 1987).

Another important topic has to do with the time scale that needs to be considered when calculating the expected number of offspring of an organism in an environment and population, that is, how many generations need to be considered in order to calculate organisms' expected number of offspring (Beatty and Finsen 1989). This is, in fact, an important problem since without agreement on the timescale to be considered there does not seem to be a univocal form of calculating fitness values, and therefore, to say whether an organism is fitter than another (but see Millstein 2016).

A different question concerns whether extrinsic factors, such as population size, should be considered besides the physical features that characterize an organism when predicting its expected number of offspring (Gillespie 1977; Sober 2001). In other words, whether fitness is just a function of the features of organisms or whether it is a function of both the features of organisms and extrinsic factors (Millstein 2016). I will come back to this problem in § 4.2.

Finally, there is also a debate concerning how fitness as an ability (i.e. non-mathematical or ecological fitness) and fitness as the expected number of offspring (i.e. mathematical or predictive fitness) can be connected. In this regard, some authors have claimed that these two faces of fitness are different but complementary notions (Brandon 1978, 1990; Millstein 2016). Yet, other authors, such as Mohan Matthen and André Ariew (2002) or Denis Walsh (2007), argue that these two notions of fitness cannot be reconciled. This is the issue I will deal with here in more detail, given that, as I will argue in section 4, my account of fitness as a causal disposition can distinguish between ecological and predictive fitness in a way that PIF is not able to do.

Ecological fitness is defined in a causal sense, insofar as it refers to the physical features of an organism that cause its survival and reproduction. Predictive fitness, conversely, refers to the organism's expected number of offspring and is defined in a mathematical form. It is on the basis of ecological fitness that biologists calculate the relative probabilities of survival and reproduction of an organism, i.e. its expected fitness. This mathematical formulation of fitness should not be confused with fitness itself. However, although within PIF fitness refers to an ability, it ends up being identified with the mathematical parameter, i.e. with the expected number of offspring of organisms (Abrams 2007). Due to this identification, there are problems in order to explain how the ecological face of fitness is connected to the predictive one: we only seem to have predictive fitness. As Roberta Millstein claims:

> It is hard to know how to reconcile fitness as a propensity, based on the heritable physical characteristics of an organism (i.e., the interpretation of fitness that addressed the explanatory circularity problem and the mismatch problem), with fitness$_1$, fitness$_2$ [...]. In particular, there seems to have been a shift from the *causes* of an

> organism's survival and reproductive success (its physical traits in relation to its environment) to the expected *outcomes* (the expected survival and reproductive success themselves) (Millstein 2016, 13).

Some authors have claimed that PIF, although does not explicitly eliminate ecological fitness, pushes it in the background (Bouchard 2011) or downplays it (Sober 2001). Frédéric Bouchard's description of the evolutionary process is a good instantiation of this view:

> If there is random variation among the traits of organisms and if some variants fortuitously confer advantages on the organisms that bear them then those organisms will live to have more offspring, which in turn will bear the advantageous traits, thereby increasing the frequency of the trait bearers (and their genes) in the population. Fitness is then explicitly described in populational terms: 'good' traits are replicated in a population so the fitter entity is the one with most descendants. In other words, fitness is usually about differential reproductive success (Bouchard 2011, 107).

The equation between ecological and predictive fitness has also been assumed when criticizing the causal role that PIF attributes to fitness. This can be seen in Sober when he argues that the viability of fitness (*as something measured* by the life expectancy of the organism) *does not cause* the organism to live a certain length of time. And, in the same way, the fertility of fitness (*as something measured by* the expected number of offspring of an organism) *does not cause* the organism to leave offspring (Sober 1993; see also Shapiro and Sober 2007).

Due to the associations between ecological and predictive fitness, different authors, such as Brandon, have remarked the importance of properly distinguishing *what fitness is*, that is, the ontological characterization of fitness, from *how fitness is measured* (Brandon 1990). However, not much work has been done regarding the ontological characterization of fitness. Contributing to the ontological clarification of the notion of fitness is the main aim of this chapter. In particular, I will attempt to offer an ontological characterization that helps to clarify what fitness is and how it relates to the expected and actual number of offspring of an organism.

I consider that some of the problems that PIF encounters in properly making sense of the causal power of fitness, as well as in reconciling ecological and predictive fitness, rest on the fact that it does not recognize the ontological reality of fitness as a disposition. In section 4, I develop a characterization of fitness as a causal disposition which, I claim, solves the difficulties of making sense of its causal power.

3. The ontology of properties: a brief introduction

In characterizing fitness as a causal disposition, I have approached it from the metaphysical framework of Mumford and Anjum's dispositional theory of causation (2010, 2011, 2013, 2018). In this section, I focus in more detail on the dispositional theory of causation (§ 3.2), after briefly introducing the more general debate over the ontology of properties in which this account is developed (§ 3.1).

3.1 Categorical and Dispositional Properties

In metaphysics, properties are generally ontologically considered to be of two different kinds, namely categorical and dispositional. Dispositional properties (e.g. solubility, fragility...) are powerful properties, i.e. powers that allow the object that possesses them to manifest a kind of behavior (manifestation) when the appropriate circumstances are given. Categorical properties are more difficult to define. Generally, they tend to be characterized as manifested qualities, that is, mere ways objects happen to be (e.g. red, spherical…). Yet, some authors, such as Alexander Bird, have considered that categorical properties are better conceived of in negative terms, that is, in contrast to what dispositional properties are. Thus, whereas dispositional properties are powerful properties, categorical properties do not necessarily require to manifest any behavior in response to a given stimulus (Bird 2007, 66-67).

The metaphysical debate regarding which kind of properties there are in the world is historically polarized into two main positions: monism and dualism. According to monism, there is only one kind of properties. For categorical monism or categoricalism, only categorical properties are ontologically real. In contrast, dispositional monism or pandispositionalism defends that there are only dispositional properties. For dualism, both dispositional and categorical properties ontologically exist.

Categoricalism

Categoricalism has its roots in the logical positivism of the early 20th century. According to the logical positivists, the ontological status of dispositions is doubtful since they cannot be directly observed. Dispositions, they claim, *are not actual properties* of objects but they refer to how an object will behave when certain circumstances are given. This non-actual character of dispositions also makes dispositional expressions problematic. For the logical positivists, sentences have meaning when the circumstances that make them true or false can be known. In this regard, it is difficult to establish the conditions of truth or falsity of dispositional expressions such as '*x* is soluble' since dispositions such as solubility cannot be directly observed.

In order to solve the problem concerning dispositional expressions, the logical empiricists equated dispositions to their manifestations. Thus, a disposition refers to the behavior an object manifests under certain circumstances. Both the circumstances for an object to manifest a certain behavior as well as the kind of behavior it manifests are empirically accessible, which means that the truth conditions of dispositional expressions can be known by means of them. The logical empiricists, therefore, analyzed dispositional expressions in terms of conditionals where both the circumstances for an object to manifest a behavior and the kind of behavior it manifests are explicitly given. In this regard, Rudolf Carnap proposed a material conditional, according to which dispositional expressions such as '*x* is soluble' are analyzed as: "for each object *x*, if *x* is put in water, it is soluble iff it dissolves" (Carnap 1936).

By means of the conditional, the truth conditions of dispositional expressions are known. Thus, the conditional concerning the solubility of *x* is true if *x* is put in water and *x* dissolves, whereas it is false if *x* is put in water and it does not dissolve. The material conditional, notwithstanding, has been charged for not being able to establish the truth conditions of ascribed dispositions to *x* when the object is not put in water (Ryle 1949). In order to avoid this problem, an alternative to the material conditional has been offered, namely a counterfactual conditional. According to this conditional, dispositional ascriptions, such as '*x* is soluble', are analyzed as "*x* would dissolve if put into water at *t*" (Ryle 1949. See also Goodman 1954; Quine 1960).

There are different counterexamples that have been offered against this conditional analysis, including the possibility of finkish dispositions (Martin

1994)[13], cases of mimickers[14] (Smith 1977; Prior, Pargetter, and Jackson 1982; Lewis 1997; Armstrong 1997), and the presence of masks (Johnston 1992) and antidotes (Bird 1998a). In face of these problems, there are authors who have opted for elaborating a better version of the conditional that could avoid them (Prior 1985; Lewis 1997; Mellor 2000; Fara 2005). Yet, others have argued that these counterexamples are not really so, and thus, it is not necessary to elaborate a more sophisticated version of the conditional (Choi 2003, 2005, 2006, 2008).

A different criticism for the logical empiricist view is based on the fact that its approach to dispositions is mainly conceptual and not ontological (Armstrong 1969). The main interest for the logical positivists is on how dispositional expressions can be put such that they can be testable. In this sense, dispositional ascriptions do not refer to an actual property of the object but to the idea that the object will manifest a particular behavior when certain

---

[13] Finkish dispositions are those that are lost or gained after the triggering conditions are given. In this kind of cases, an object that is said to possess a disposition will lose it once the triggering conditions are given. In the same way, an object that does not possess a disposition is able to gain it once the triggering conditions are given. The idea behind these cases is that the circumstances for losing or gaining a disposition are the same as those for triggering a disposition. Charlie Martin (1994) illustrates finkish dispositions with the following example: let's consider a dead wire that is connected to an electro-fink, that is, a device that senses when the wire is about to be touched by a conductor, and which make the wire live (that is, able to conduct electricity). In every such circumstance, if the wire were touched by a conductor, then it would become live and able to conduct electricity. In this case, a dead wire, which is not able to conduct electricity, gains this disposition once the triggering conditions, i.e. being touched by a conductor, are given. In the same line, Martin also considers a reverse device in which a live wire is connected. In this case, when the wire is about to be touched by a conductor, the wire gets dead such that it is not able to conduct electricity. Thus, a live wire, which is able to conduct electricity, loses its disposition once the triggering conditions are given.

[14] Cases of mimickers are those in which a particular object seems to manifest a disposition, although it does not really possess that disposition. An example of a mimicker could be an iron cooking pot that is attached to a bomb with a very sensitive detonator that can make it shatter (Bird 2007). In this case, the pot is mimicking the disposition of fragility since the manifestation is the same one associated with it. Yet, the pot does not really have this disposition.

circumstances are given. However, as David Armstrong argues, there must be something actual in the object that explains the behavior it manifests under certain circumstances. According to Armstrong (1969), this actual element refers to the categorical properties of the object. Categorical properties (together with the laws of nature) are responsible for the object to manifest a certain behavior when certain circumstances are given. This is the view called categorical monism or categoricalism (Mumford 1998).

Categorical properties, unlike dispositions, are powerless. Yet, they can manifest a dispositional character: categorical properties are also considered to be able to perform a causal power under some circumstances. This causal power, notwithstanding, is not essential to categorical properties, but they obtain it as a consequence of the laws of nature.[15] Alexander Bird introduces this idea in the following terms:

> The categoricalist regards that [dispositional] character as being imposed upon a property by the laws of nature. Let P be a property that as a matter of fact has the dispositional character given by $D_{(S,M)}$. The categoricalist explanation of this is that there is a law which says that whenever something is P and also S, then it is also M. For example, the property of being negatively charged has this character: objects that are negatively charged have the disposition to attract positively charged objects. On the categoricalist view this does not reflect any essence of (negative) charge, but holds instead because a law of nature relates positive and negative charge (or more accurately governs charge in such a way that the force between opposite charges is attractive) (Bird 2007, 68).

Although, as previously stated, for categoricalists, categorical properties are the only ontological properties that exist, they accept a

---

[15] There are two main views about laws that are compatible with categoricalism: the regularity view (Lewis 1973), and the nomic-necessitation view (Armstrong 1983). It is not within the scope of this thesis to develop these accounts of laws in detail. But see Bird (2007) for an overview of them.

conceptual distinction between categories and dispositions. Dispositional ascriptions are made in those cases in which we do not know yet which categorical properties of an object are responsible for the kind of behavior it manifests under certain circumstances. John Leslie Mackie puts this idea as follows:

> […] failure to distinguish between questions of epistemology and ontology may encourage the rationalist view. Not only in every day knowledge but also in scientific knowledge we have to deal largely with properties introduced in the dispositional style. The layman knows of fragility or conductivity only either as the causal property of acting or reacting in a certain way or as some otherwise unspecified categorical feature by whose possession a thing is causally apt for so acting or reacting (Mackie 1977, 368).

Pandispositionalism

Other authors have also considered the same criticism as Armstrong against the logical empiricist view, namely there must be something actual in the object that explains the behavior it manifests under certain circumstances. This actual element, nonetheless, refers to dispositional and not to categorical properties. David Hugh Mellor, for instance, has argued in this line (1974). Mellor has criticized categoricalism on the basis that, in order for it to be true, it would be necessary to see the kind of behavior the object manifests under some circumstances before associating its categorical basis with a particular manifestation. However, Mellor argues, it seems that we are able to attribute dispositions to objects before they manifest a behavior. This is the case, Mellor claims, of nuclear plants, since we know that some materials are disposed to explode without the necessity of seeing them explode. According to Mellor, therefore, the ontological aspect of the object that allows us to attribute it a kind of manifestation under certain circumstances is the dispositional property it has.

Dispositional properties, Mellor argues, are actual properties. What is in potency is not the disposition but the manifestation of the disposition (Mellor 1974; see also Mumford 1998; Heil 2005). Dispositional properties are actually given in the object regardless of whether they are manifested or not. In this regard, what establishes the difference between a glass that remains brittle at all times it remains unstruck and a glass that loses it brittleness at all times it remains unstruck, for instance, is that in the first case the glass has the disposition to break whereas in the second it has not. This idea is also illustrated by Mumford when claiming that:

> Arguably, it makes sense only because we conceive of dispositional properties as real properties that are there whether or not they manifest. Thus, there is some difference in the world between a wire being live and being dead. This is a difference which [...] depends on no test or manifestation. *Having or not having the disposition is itself the difference* (Mumford 1998, 55. Emphasis added).

Within this framework, dispositional properties are causal properties. They are the ones responsible for the behavior objects manifest under certain circumstances. Dispositions, therefore, are equated to powers. This position is called *dispositional monism* or *pandispositionalism* (Mellor 1974; Shoemaker 1990; Bird 2007; Mumford 2004; Mumford and Anjum 2011). Generally, dispositionalists consider dispositions to be characterized by the following features (Molnar 2003): (1) *directedness* (or tendency) -dispositions tend towards their manifestations without necessitating them. Once a disposition is triggered, the manifestation might not be produced due to some factors that may interfere with the disposition itself; (2) *Independence* -dispositions do not need their manifestations to exist; (3) *actuality* -dispositional properties are as actual as categorical properties are said to be for categoricalists. What is in potency is the manifestation of the disposition (Mumford 1998; Molnar 2003; Heil 2005); (4) *objectivity* -dispositions are not anthropocentric properties. They do not depend on human construction, thought or perception to exist; and (5) *intrinsicality* -the dispositional properties that an object has do not depend on which other objects exist (Molnar 2003).[16]

Pandispositionalist have offered different arguments against categoricalism. One of them is the difficulty for this account to determine the identity of properties. According to categoricalists, the identity of properties is a *brute fact.* It is given by means of their essences. This position is called 'quidditism' (Bird 2007) and is regarded as problematic because essences cannot be observed. Pandispositionalists claim that it is easier to determine the identity of properties when they are conceived of as dispositions. Since dispositions are powerful properties, it is the type of manifestation towards which a disposition tends that determines its identity or allows to individuate it. Manifestations, unlike essences, can be observed, and thus, they allow to determine the identity of properties in a better way (Bird 2007, 78).

Another difficulty for categoricalism relies on the way it explains the causal power of properties, namely by means of the laws of nature. Pandispositionalists have emphasized the contingency of the laws of nature: there could be different possible worlds with different laws. Thus, a particular categorical property might have different causal powers in a world with different laws (Shoemaker 1980, 1984; Bird 2007). The problem with this position is that despite having different causal powers due to the different

---

[16] The intrinsicality of dispositional properties has been questioned by authors such as McKitrick (2003) who considers that there can be extrinsic as well as intrinsic dispositions.

laws that govern the different worlds, the categorical property will remain the same since "the quidditist conception of properties is that they have primitive identity, identity that is completely independent of their powers" (Bird 2007, 74).

But pandispositionalism has also been challenged. One of the criticisms highlights its difficulties to establish the identity of properties. Being dispositional, properties are epistemically inaccessible, and thus, their identity cannot be established. The so-called epistemic (Swinburne 1980) and individuative (Lowe 2006; Bird 2007; see also Marmodoro 2010) regresses illustrate this idea. Both kinds of regress are based on the idea that, in a pandispositionalist ontology, the manifestation of a disposition must be another disposition. Both the identity and the epistemic access to the disposition will only take place through the manifestation of the disposition. Yet, since this manifestation will be another disposition, we end up in an epistemic and identity regress (Swinburne 1980; Lowe 2006). Within this kind of ontology, therefore, we are not able to epistemically access nor to know the identity of any property.

Besides the previous problem, the possibility of multi-track dispositions also poses a difficulty for pandispositionalism to establish the identity of properties. Multi-track dispositions are those that have more than one kind of stimulus and manifestation (Ryle 1949; Bird 2007). Elasticity might be an example of a multi-track disposition. It can be claimed that something is elastic when it is able to recover its original shape when we stop bending it, or when it expands until its original size when we stop pressing it. In these cases, we can properly talk of the objects as being elastic, although both the stimuli (bending and pressing) and the manifestations (to bend and to compress) are different in each situation. Multi-track dispositions are problematic since it is not clear whether there is one and the same dispositional property that is manifested in different ways when different stimuli are given or, conversely, whether there are different dispositions with different stimuli and manifestations that are referred to with the same term (Lowe 2010).

Dualism

According to authors such as Edward Jonathan Lowe and Richard Swinburne, in order to have epistemic access to dispositional properties and be able to fix their identity, their manifestation must be different in kind from the dispositional properties themselves, i.e. they need to be categorical (Lowe 2010, 22). Thus, we need both dispositional as well as categorical properties in our ontology.

This approach according to which both categorical and dispositional properties ontologically exist is called *dualism*. In this case, the distinction between categories and dispositions is both ontological and conceptual, that is, it is not that there are two different ways of talking about one kind of ontological property (as per categoricalism), but that there is a division in reality such that there are two kinds of ontological properties we can talk about (Ellis 2010).

Dualists have offered different arguments for the distinction between dispositional and categorical properties (Mumford 1998). Firstly, categorical properties are the bases of dispositions. Categorical properties explain dispositions, and something cannot be identical with that that explains it. Secondly, the same kind of disposition can have different ontological bases. Therefore, a kind of disposition is not identical with a particular categorical basis. Finally, if the categorical basis would be identical with a disposition, then by knowing the categorical basis of an object we would be able to know its dispositions and capacities, but this is not the case.

Besides avoiding the individuative and epistemic regress, dualists have also argued that categorical properties are, in fact, necessary for locating both dispositional properties and their manifestations. Without categorical properties, we would not be able to know where dispositional properties are. Location, Ellis (2010) claims, is a categorical property.

Despite these advantages, dualism is not exempt from problems. One of them refers to the fact that dualists do not offer accurate criteria to properly distinguish categorical from dispositional properties (Mellor 1974; Choi 2008). Another difficulty for dualism refers to its incapacity to properly establish the kind of relationship that exists between categorical and dispositional properties. Some dualists have claimed that the relationship is one of grounding (Lowe 2010), whereas others have considered it to be a causal one (Prior, Pargetter, and Jackson 1982). In this regard, categorical properties are considered to be the causal bases for dispositional properties, that is, they are responsible for the causal power that is attributed to the disposition. Elizabeth Prior, Robert Pargetter, and Frank Jackson present this idea as follows:

> There is, that is, a reason why the glass is fragile. This reason involves a causally relevant property (or property complex) of the glass, which we will call the *causal basis* of the disposition. [...] For each disposition we can specify a pair of antecedent circumstances and manifestation which together determine the disposition under discussion. In the case of fragility, the pair is (roughly) <knocking, breaking>, in the case of water solubility the pair is <putting in water, dissolving> and so on for the other familiar cases. By "a causal basis" we mean the property or property-complex of the object that, together with the first member of the pair -the antecedent circumstances- is the causally operative sufficient condition for the manifestation in the case of 'surefire' dispositions,

> and in the case of probabilistic dispositions is causally sufficient for the relevant chance of the manifestation (Prior, Pargetter, and Jackson 1982, 251).

Yet, the causal interpretation of the relation between categorical and dispositional properties has also been questioned. George Molnar has claimed that there are some cases in which it is problematic to explain from where the dispositional property obtains its causal power. On the one hand, it seems that some dispositional properties are ungrounded in the sense they do not seem to have categorical bases. On the other hand, there is nothing that precludes the basis of a particular dispositional property to be another disposition (Molnar 1999).

Dualism has also been charged for being a mere form of epiphenomenalism. Insofar as dualist accounts consider that the causal power attributed to a dispositional property is really given in its categorical basis, dispositional properties are causally impotent, and thus, they are mere epiphenomena, that is, they might have ontological reality although they do not have causal autonomy (I will say more on this in Chapter 3). If dispositional properties are not considered to have causal powers on their own, then it is not clear why they should be characterized as such (Mumford 1998, 114-115).

Epiphenomenalism or the lack of causal powers in dispositions is, in fact, one of the main charges against dualism. It favors the idea that it does not seem to be really necessary to postulate two kinds of properties, i.e. categorical and dispositional, since both categoricalism and dispositionalism are able to explain everything that dualism can with more ontological parsimony.[17] The debate concerning the ontology of properties, therefore, seems to be better addressed from the categoricalist and pandispositionalist framework. Despite its problems and criticisms, some pandispositionalists have argued that the advantages of pandispositionalism outweigh its disadvantages (Molnar 2003; Hüttemann 2007; Chakravartty 2013b, 2017). In particular, it is argued that a dispositional ontology allows for a better understanding and clarification of metaphysical questions such as those of modality (Molnar 2003), laws of nature (Mumford 2004; Bird forthcoming), or causation (Molnar 2003; Mumford and Anjum 2011).

For the particular purpose of this chapter, namely to clarify the concept of fitness from a metaphysical perspective, the causal dispositionalism proposed by Mumford and Anjum is particularly useful. As I will argue in section 4, the conceptual resources that causal dispositionalism offers (e.g. threshold, combination of powers, resultant effect), constitute a promising metaphysical framework for understanding the notion of fitness as a disposition of organisms. Biological phenomena show distinctive traits that the physics-oriented debate on dispositions has not taken into account. I

---

[17] In this regard, it is important to notice that there are two accounts that also try to argue for both dispositional and categorical properties, but that are not dualist accounts. I refer here to neutral monism (Mumford 1998) and the identity theory (Martin 1997; Heil 2005, 2010). According to *neutral monism*, there is only one kind of ontological properties. Yet, we are not able to know whether this kind of properties is categorical or dispositional, and, as a consequence, we cannot favor dispositionalism over categoricalism or vice versa (Mumford 1998). For neutral monism, therefore, both categorical and dispositional terms are equally valid, insofar as they refer to two different forms of talking about one and the same ontological property (Mumford 1998, 190). According to *the identity theory*, properties are both categorical and dispositional at the same time. In this sense, every property is a manifested quality and also a power, so that properties are better conceived of as 'powerful qualities' (Heil 2005). Concerning neutral monism, Mumford ends up considering that the ontological status of properties is, in fact, dispositional (Mumford 2004). Regarding the identity theory, it has been questioned on the basis that it has problems in explaining how the categorical and the dispositional aspects of a property are different and identical at the same time (Engelhard 2010).

consider that the complexity of biological entities, or the emergent character of biological organization, require to re-examine the notion of disposition when applied to biological phenomena and, particularly, to the case of fitness. In this sense, causal dispositionalism is particularly useful for the purpose of characterizing fitness because it is compatible with the specific features of biological phenomena, it does not imply physical reductionism, and its role in clarifying the ontological status of other biological concepts such as that of 'genes' has already been defended:

> An adequate understanding of what, for instance, genes do in biology requires acknowledgment of complexity and context-sensitivity, two key notions that are emphasized in the dispositional account. A crucial part of the understanding of what genes can and cannot do, we will argue, is to accept that they operate in a dispositional way, and we thus see biology as exemplifying the general nature of causation as here developed (Mumford and Anjum 2011, 2018).

In the following section, I will present the causal dispositionalism proposed by Mumford and Anjum in more detail.

3.2. The dispositional theory of causation

Within the pandispositionalist framework, Mumford and Anjum (2011) have developed a dispositional theory of causation (hereafter DTC). According to this account, "[c]ausation [...] is about properties and, more specifically, the powers or dispositions that make up these properties" (Mumford and Anjum 2011, 22). Dispositional properties are responsible for performing causal work insofar as they are powerful. And, since dispositions are properties of things, there is a *derivational* sense in which causation might involve objects, facts, and events: all of them are entities involved in causation due, precisely, to the properties they contain (Mumford and Anjum 2011).

For DTC, dispositional properties, insofar as powerful, are the ones responsible for producing their own manifestation, and production "is clearly itself a causal notion" (Mumford and Anjum 2011, 8). According to DTC, causation is not a relation between events, but the *process* that takes place once dispositions have been triggered and their manifestations are being produced (Mumford and Anjum 2011, 6, 209). In a causal process both causes (i.e. triggered dispositions) and effects (i.e. the manifestations associated with dispositions) are *simultaneous,* that is, they do entirely coincide temporally during the whole time the causal process lasts. For DTC, therefore, there is not temporal priority of cause over the effect. Furthermore, there is no cause without effect, nor effect without cause since "[t]he cause takes time to do its work, but the effect coincides with that time entirely" (Mumford and Ajum 2011, 124). Besides being a process, causation is also primitive. It cannot be reduced to something else (Mumford and Anjum 2011, 158). As Mumford and Anjum make explicit, it is not that causation can be explained in terms of dispositions, but rather that causation is a process that involves dispositions. DTC, therefore, is not an analysis of causation in terms of something else, but a characterization of what causation ontologically is.

Mumford and Anjum consider this primitive aspect of causation as one that allows them to differentiate their position from Humeans accounts of causation.[18] According to David Hume, causation cannot be directly perceived. What we, humans, perceive is an event that is followed by another event. Yet, the connection between them, what would be the real 'cause', is

---

[18] As we will shortly see, the authors also differentiate their account of causation from the anti-Humean view of causation.

not perceived. Hume, therefore, concludes that the idea of causation is *inferred* from other, more simple ideas that are directly perceived, such as spatial contiguity, temporal priority, and constant conjunction (Mumford and Anjum 2011, 199).

Besides arguing, against Humean accounts of causation, that causation is primitive, Mumford and Anjum also questioned Hume's idea that causation cannot be empirically perceived. In their view, causation can be directly known from experience insofar as to perceive causation means to perceive powers at work. And we are able to perceive powers throughout bodily sensations insofar as we are physically embodied (Mumford and Anjum 2011, 201). It is with our bodies that we are causal agents, both passively, insofar as we are able to be acted upon (i.e. to perceive a cause acting on us such as when the wind pushes us), and actively, insofar as we are able to act upon things (i.e. to cause an effect such as lift, push, pull…). For Mumford and Anjum, the sense that allows us to be aware of causal agency is *proprioception*, which refers to "the sense of a required effort" (Mumford and Anjum 2011, 207-208). It is precisely by means of proprioception that we can properly respond to the causes that act upon us (by considering, for instance, the kind of effort we must offer in order to overcome them), as well as to manage the way we act upon things (by means, for instance, of being aware of the kind of resistance that we are met when acting on something) (Mumford and Anjum 2011, 207-208).

One of the virtues of DTC over other accounts is that it makes explicit the complex character of causation. According to Mumford and Anjum, causation is more complex than assumed. In particular, they are against those accounts, such as the counterfactual one, that conceives causation as a relation between *a* cause and an effect (see also Lowe 2002). In this kind of accounts, *a* cause is recognized from a set of background conditions for producing a certain effect. Within this framework, the cause is the element that is added to or subtracted from the background conditions such that a certain effect is produced.

Mumford and Anjum illustrate this distinction between causes and background conditions with an example. Let's consider two different scenarios (*S1* and *S2*) wherein we want to know the cause of a match lighting. In *S1*, we are in a room wherein there is oxygen and the matchbox is rigid and dry. In *S2*, we are in a room wherein there is no oxygen and the matchbox is also rigid and dry. In both scenarios, in order to light the match, we strike it against the matchbox. Yet, only in *S1* but not in *S2* (due to the absence of oxygen) the match would light. In *S1*, the presence of oxygen in the atmosphere, as well as the rigidity and dryness of the matchbox would be considered the background conditions. It is the strike of the match against the matchbox what causes it to light. However, in *S2*, no change has been produced despite the striking. Let's also imagine that in order to light the match in *S2*, we start to continuously strike it against the matchbox until some oxygen enters the room due to some failure and the match lights. In this case, the striking of the match against the matchbox, as well as the rigidity and dryness of the matchbox would be considered as the background conditions of the situation, while the presence of oxygen, as the *difference* that makes possible for the match to light, would be considered to be the cause of the event.

In both scenarios, and due to the distinction between causes and background conditions, it is possible to recognize only one element as a cause of an effect, i.e. the lighting of the match. However, DTC does not accept the distinction between causes and background conditions. Within this framework, the distinction is considered to be an epistemological rather than an ontological one (Mumford and Anjum 2011, 32; see also Austin 2016a). For DTC, all powers must be considered when looking for the cause of an effect since all of them are acting together in its production. Yet, for pragmatic reasons, we might decide to focus on just one of them.

The causal relevance of multiple dispositions contributing to a causal process plays a crucial role in Mumford and Anjum's understanding of the complexity of causation. In their view, dispositions do not work in isolation. In fact, it is really rare to find a disposition that acts alone (Mumford and Anjum 2011, 27. See also Molnar 2003). Dispositions combine, i.e. they come together to work jointly (Mumford and Anjum 2011, 100), in order to produce a certain effect. The idea behind composition is that the same power can produce different effects depending on what other powers it combines with (Mumford and Anjum 2011, 35). Molnar calls 'pleiotropy'[19] this idea that a particular disposition can participate in the production of different effects (Molnar 2003). It is important to point out here that effects must not be equated to manifestations within DTC (although see McKitrick 2010 for a rejection of this distinction). The manifestation of a disposition is its contribution to an effect (Molnar 2003; see also Austin 2016a). A type of disposition does always have the same kind of manifestation. Yet, depending on what other dispositions a particular disposition is combined with, the effect might vary. In other words, the manifestation of a disposition is the end-state of the function that defines the property, exhibited whenever the disposition is triggered, while the effect of a disposition is the resulting,

---

[19] This is, in fact, a biological term that refers to the idea that a particular gene might affect or intervene in the production of different phenotypes.

particular event or property to which the manifestation contributes once the disposition is triggered (Nuño de la Rosa and Villegas manuscript).

According to *'compositional pluralism'* (Mumford and Anjum 2011, 86), the combination of dispositions might take place in different forms which, recovering a classic metaphysical distinction, can be classified into two main ways: linear and non-linear (Mumford and Anjum 2011, Ch. 4). When compositions are linear, the result that is obtained from the combination of different powers is the same as the result that is obtained from the sum or addition of the effect of each power acting in isolation (Mumford and Anjum 2011, 88).

An example of linear composition might be a billiard ball on which two forces (e.g. the north force and the west force) are acting together such that the ball moves in a north-west direction. In this case, the ball would have move in the same north-west direction if one of the forces would have acted before the other and not at the same time. In many natural phenomena, notwithstanding, this kind of interaction is not the one that is given, insofar as the components that interact give rise to an effect that is not the product of their addition or subtraction. Non-linear compositions, therefore, do not obey the principle of additivity. Overdose, escalatory, and antipathetic cases are some examples that illustrate the non-linear combination of dispositions.

Overdose cases illustrate non-linearity insofar as there is not a proportion between the number of powers of the same type that are added and the resultant that is produced. In these cases, it is possible for a disposition *D* to manifest *R*. Yet, to add more dispositions of type *D* does not imply that the resultant *R* would be manifested with a higher intensity. In fact, it is possible for the addition to produce a resultant that is completely different from *R*. For instance, eating chocolate might cause us to feel pleasure. Yet, it is not that the more chocolate we eat the more pleasure we feel. In fact, we might feel sick after eating a certain number of chocolate bars. Mumford and Anjum also highlight that overdose cases are not only produced by adding more powers of the same type, but they might also be produced by letting a given power to act during a long period of time. Ice, for instance, works as an anti-inflammatory but only when it is applied for a period of five minutes. When it is left longer than that, it produces precisely the opposite effect by leading to inflammation (Mumford and Anjum 2011, 89).

Escalatory cases refer to those cases in which a small power added to a causal process can produce a big effect. This is the case of chaotic systems, as illustrated by the so-called 'butterfly effect', where the effect is bigger than the sum of the component powers (Mumford and Anjum 2011, 90). Finally, *antipathetic cases* are those cases, well known in medicine, wherein two individual powers can separately dispose towards a given effect, although they both dispose away from it when working together. Thus, lonidine and beta-blockers separately lower blood pressure, but taken in combination have the opposite effect of raising blood pressure (Mumford and Anjum 2011, 91).

The consideration of non-linear combinations within DTC illustrates the complex way causation might take place. Mumford and Anjum characterize non-linear combinations as 'emergent causal interactions' (Mumford and Anjum 2011, 96). In their view, non-linear composition "suggests some variety of emergentism because two powers could produce a novel phenomenon that can be both surprising (weak emergence) but also genuinely productive of something new in the world (strong emergence)" (Mumford and Anjum 2011, 87-88). Mumford and Anjum do not really go deep regarding the question of emergence. Yet, they seem to be committed to the idea that weak emergence corresponds to epistemological emergence, since the relevant aspect of the higher-level, emergent property, is not its novel character but that it is, in some sense, unexpected. The phenomenon is epistemologically surprising since it cannot be expected from the powers at work. In the same line, they also seem to associate strong emergence with ontological emergence, insofar as it refers to the introduction of something ontologically new in the world. As we will see in more detail in the next Chapter, the equation of ontological and epistemological emergence to strong and weak emergence, respectively, has been questioned. Thus, although it is important to highlight that emergentism has a place within the metaphysical framework of DTC, insofar as it recognizes that dispositions that result from a causal process might be emergent, this is not the way I will approach emergentism in this thesis.

The consideration of non-linear combinations within DTC as well as of the idea that causation is not a relation between events allows Mumford and Anjum to claim for the inadequacy of the classic model used to represent causation, i.e., the neuron model. The neuron model of causation represents the relata of causation (generally considered to be events) by means of circles, and the causal connections between them by means of arrows. When a certain event *a* is 'fired' and is causally connected to another event *b*, then event *b* occurs, as long as there is not any inhibitor that avoids the manifestation of *b* (Mumford and Anjum 2011, 20). The neuron model, Mumford and Anjum claim, is not suitable to represent causation since, on the one hand, it conceives causation as a relation between *a* cause and the effect; and, on the other hand, it illustrates causation as a necessary relation such that given the cause, the effect must be produced if there are not inhibitors between the events. Causation, therefore, seems to be an all or nothing relation, something that Mumford and Anjum reject:

> [In the neuron model] causation is also depicted as all or nothing. When the cause occurs, and there is no inhibitor, the effect must occur. Simulatory connections guarantee the effect. They do not, in their unmodified form, show cases where there is only a certain probability of an effect occurring, or a case where an effect can occur to some degree, depending on the strength of the cause. On the other side of the coin, inhibitors guarantee that an effect does not occur, regardless of what else is going on. Again, there is no allowance for something that might inhibit an effect to some degree, or lower the probability of an effect occurring (Mumford and Anjum 2011, 21).

As an alternative to the neuron model, Mumford and Anjum propose the vectorial one (see Fig.3). In the vector model, dispositions are represented by means of vectors, which have two main features: intensity, illustrated in the model by the length of the vectors; and directionality, marked by where the arrowheads are pointing to.[20] Causation occurs when powers have accumulated so that they are enough to trigger a certain effect. There are frequently many powers involved in the production of an effect, each disposing in the same direction, in varying degrees. Those dispositions that dispose towards a particular effect are the causes of the effect. Yet, there might also be dispositions that dispose away from the effect. These are what Mumford and Anjum call 'countervailing factors'. Countervailing factors affect how an effect is produced insofar as they intervene with and affect the rest of dispositions that dispose towards the effect. In this regard, countervailing factors could also be considered as causes. As Mumford and Anjum make clear, to consider countervailing factors as causes or not is a

---

[20] As Mumford and Anjum have made explicit, the representation in the vector model is made in one-dimensional-quality space. Yet, this is so for questions of simplicity. In fact, the vector model allows for multi-dimensional quality space (see Mumford and Anjum 2011, 45). Throughout the thesis, and in order to guarantee simplicity, I will also use the one-dimensional quality space to represent the dispositional character of fitness.

terminological question. However, they prefer to reserve the term 'cause' only for those dispositions that dispose towards the effect (Mumford and Anjum 2011, 33). I will maintain this conceptual distinction between causes and countervailing factors in this thesis as well.

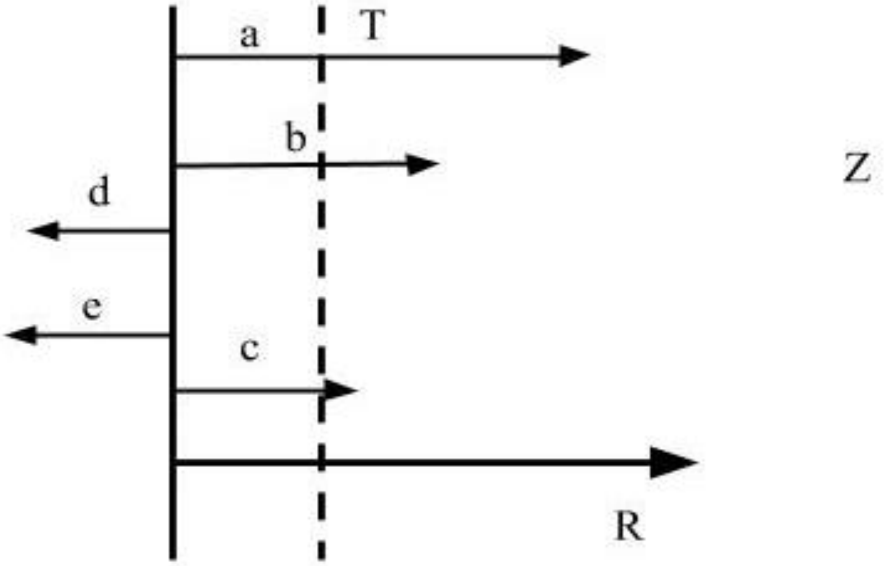

Figure 3. **Dispositional Theory of Causation.** Vectors *a*, *b,* and *c* refer to different triggered dispositions (i.e. causes) that are combined and oriented towards a certain full effect *Z.* Vectors *d* and *e* refer to countervailing factors. They are dispositions that participate in the causal process and affect how the resultant is manifested. In this case, they avoid the manifestation of the full effect *Z*. *T* indicates the threshold that has been reached such that some effect, although not the full effect *Z*, is produced. This effect that results from the combination of the triggered dispositions is the resultant *R*. Due to its simultaneous character, both causes (vectors *a*, *b* and *c*) and countervailing factors (vectors *d* and *e*), and the resultant *R*, are simultaneously given during the whole time the causal process lasts. Vectors have two main features: intensity, that is illustrated by the length of the arrow, and direction, illustrated by the way the arrowhead is pointing to.

The effect that results from the combination of dispositions is what Mumford and Anjum refer to as the 'resultant' (see Fig.3, R). Due to its simultaneous character, insofar as there is a cause there is always an effect taking place. Even in those cases in which there is an equilibrium state regarding both causes and countervailing factors (such as in a tug-of-war where the two teams are equally balanced and neither of them gains an advantage (Mumford and Anjum 2011, 30)), there is also an effect. This effect will be represented in the vector model as a 'zero resultant vector' (see Fig.4). The relevant aspect of these cases is to recognize that there is a causal process taking place, and not to confuse them with a situation in which there is not causation, i.e. when there are no dispositions acting.

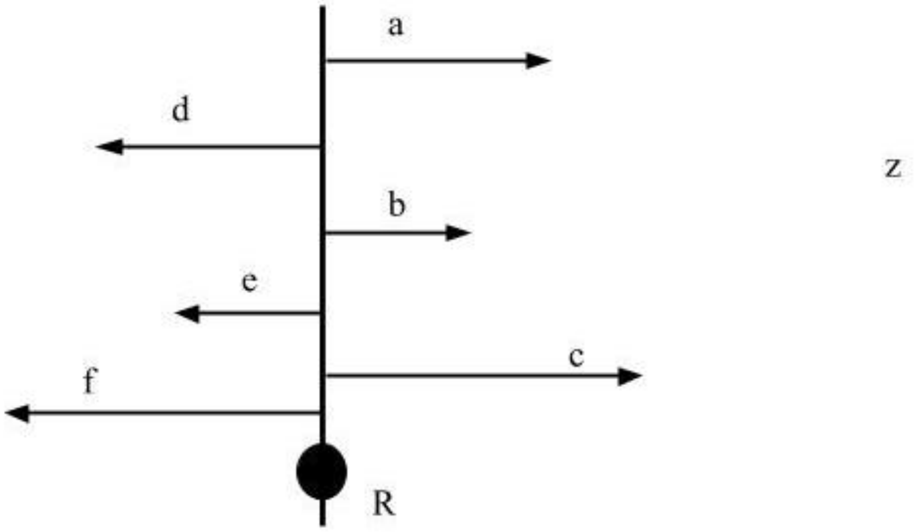

Figure 4. A zero resultant vector. In this causal process, since both causes (i.e. vectors *a*, *b* and *c*), and countervailing factors (i.e. vectors, *d*, *e*, *f*) have the same intensity, the resultant of the causal process refers to a state of equilibrium of the process itself. The resultant does not dispose either towards the effect *Z* or away from it. (Figure reproduced from Mumford and Anjum's 2011, 29).

As it has been noticed above when criticizing the neuron model, DTC rejects that causes and effects are connected by a necessary relation, as per some anti-Humeans accounts of causation. Mumford and Anjum offer different arguments against necessitarianism, the main idea being that the necessity of causation is often assumed rather than argued. One of the arguments they use to illustrate this point is the 'antecedent strengthening' one. This argument is based on the idea that necessity would support monotonic reasoning, according to which, where *A* necessitates *B*, then whenever *A* is the case, *B* is the case, such that *A* is then a sufficient condition for *B* (Mumford and Anjum 2011, 57). If there is a necessary connection between causes and effects, then it is not relevant what other events take place besides the causal one. Insofar as the cause is given, the effect is given too: "where *A* necessitates *B*, no new information or extra premise can prevent *B* if *A* is the case. If *A* necessitates *B*, then even *A* and ¬*B* necessitates *B*, according to the standard logical theory" (Mumford and Anjum 2011, 57). Thus, if it is claimed that the strike of the match is the cause of the match lighting, such that there is a necessary connection between the strike and the lighting, then it needs also to be claimed that the strike of the match together with the presence of a gust of wind, would also cause the lighting of the match. However, this seems to be difficult to accept since, at least intuitively, we can conceive that the match would not light due precisely to the gust of wind. According to Mumford and Anjum, insofar as it does not support the antecedent strengthening, causation is non-monotonic. The connection between causes and effects is, therefore, not necessary, and thus, the kind of modality that characterizes causation is not necessity. As we will shortly see, it is not pure contingency or possibility either.

For Mumford and Anjum, the kind of modality that is involved in causation is a primitive one that they call 'dispositional modality'. Dispositional modality refers to the idea that the connection between causes and effects is purely dispositional insofar as causes only dispose towards their effects. The connection between causes and effects, therefore, is not necessary. Nor is it purely contingent. Dispositional modality is a primitive, *sui generis* form of modality that is characterized for being more than pure possibility but less than necessity. It is *less than necessity,* since different interfering and preventing factors might intervene in a causal process, altering the effect or even preventing it. It is *more than pure possibility* or contingency, insofar as there is not a tendency or disposition for everything that might be (logically) possible. As Mumford and Anjum claim:

> In the sense of pure, logical possibility, we can say things such as that it is possible that the struck match turns into a chicken, that lead could turn into a gold, or that pigs could fly. But there are not corresponding dispositions to do so (Mumford and Anjum 2011, 181).

Logical possibility, therefore, is not the relevant one when talking about dispositions, but natural possibility is (Mumford and Anjum 2011, 182). In this regard, dispositional modality is more than pure possibility, insofar as it selects only a subset from a realm of possibilities. Dispositional modality favors some possibilities over others. And the possibilities that are favored are those towards which there is a natural disposition or tendency (Mumford and Anjum 2011, 189). According to Mumford and Anjum, this tendency of dispositional modality towards what it is natural illustrates the idea that natural possibility is grounded in dispositional modality, insofar as what is naturally possible is what is disposed to happen (Mumford and Anjum 2011, 182).

In order to properly represent, within the vector model, that there is not a necessary connection between the cause and the effect, Mumford and

Anjum turn to the notion of *threshold* to illustrate that a particular effect has, in fact, been produced: "the threshold is not some extra existent. It simply marks a point, the reaching of which we might identify as an effect" (Mumford and Anjum 2011, 73). The threshold can be used to mark either the full effect of a causal process, that refers to the complete manifestation of a triggered disposition, or another effect that is produced by the causal process but that it is not the full effect (see Fig.3). In these last cases, interfering and preventing factors intervene in the causal process such that the full effect is not reached. If interfered, the effect is different from the one that would have been produced without the interfering factor. If prevented, the causal process stops and no effect is being produced.

The distinction between the full effect and other effects that are not the full one might be illustrated by means of an example. Let's consider two dispositions, namely the disposition of a cube of salt to dissolve in water and the disposition of water to dissolve salt. When both dispositions are triggered, that is, once the salt is introduced in water, a causal process starts in which salt is dissolved. The *complete* dissolution of a cube of salt in water refers to the full effect of the causal process that consists in the dissolution of salt in water. Yet, one could introduce a cube of salt in water, such that a causal process starts, but picking it up before it fully dissolves. In this case, some effect (although not the full one) has been produced in the process, since a portion of the cube of salt, although minimal, has been dissolved. Since in DTC causation is a process that takes time, there is no difficulty in acknowledging that different effects, that are not the full one, might be recognized during the causal process (e.g. the cube of salt being 1% dissolved, the cube of salt being 2% dissolved... and so on and so forth) (Mumford and Anjum 2011, 123).

According to Mumford and Anjum, the way of conceiving causation in DTC allows to solve or clarify some classic metaphysical problems such as the possibility of causation by absence, the problem of induction, and the problem of the overdetermination of causation. Here, I will only focus on the last one since the form how overdetermination is dealt with in DTC will be relevant when arguing for the causal autonomy of fitness (Chapter 3, § 3.2). As we will see in the following chapter, one of the objections to the idea that emergent, higher-level properties have causal powers is that it would give rise to cases of overdetermination.

The overdetermination of causation refers to the idea that a given effect can have more than one set of independently sufficient causes (Bunzl 1979), such as when two rocks shatter a window at the same time and the window breaks, or when "an assassin's shot *c*, causes the victim's death *e*, but a second assassin also fires at the same time as the first such that this shot *d*, also causes *e*" (a classic example in the literature known as the fire-squad) (Lowe 2002, 179). This form of overdetermination is called 'actual overdetermination' (Lowe 2002, 179). Although there are other cases of causal overdetermination, such as pre-emption[21] and fail-safe cases[22] (Lowe 2002, 179), I will only consider actual overdetermination here, given that this is the kind of overdetermination to which opponents of emergentism appeal (Kim 2006). As we will see, opponents of emergentism claim that emergent properties inherit their causal powers from the lower-level properties of the parts of the system

---

[21] Causal pre-emption occurs when there is a cause *c*, which causes an effect *e*, yet, another cause *d* also takes place such that, although *d* is not the real cause of *e*, if *c* had not occurred, then *d* would still have occurred and would then have caused *e*. In cases of pre-emption, both *c* and *d* take place. However, *d* would only produce the effect *e*, and therefore, would only be the cause of *e*, in case that *c* fails.

[22] A fail-safe case of causal overdetermination is given when there is a cause *c* to produce an effect *e*, yet, in case *c* would not occur, then a different cause *d* would appear such that it causes *e*. In these cases, *d* is only given just in case *c* is not.

upon which they depend. In this regard, the effect that is caused by the emergent property is overdetermined insofar as it is also caused by the lower-level properties of the parts of the system.

For DTC, the problem of overdetermination arises regarding both the causal power that characterizes the component dispositions of a causal process and the causal power that characterizes the disposition that results from that causal process. If both the resultant and component dispositions of a causal process tend towards the same effect, i.e., they have the same causal power, then there is a case of overdetermination. Conversely, if the resultant disposition tends towards an effect that is different from the one its dispositional components tend to, then no overdetermination is given. In the first case, since the kind of manifestation towards which both the resultant and component dispositions tend is the same, both are, in fact, the same disposition (Molnar 2003, 145). In the second case, however, there are two different kind of dispositions (the resultant and its component dispositions) that tend towards different manifestations.

For Mumford and Anjum, these cases of overdetermination in which both causes and the resultant effect tend towards the same manifestation could be given. Yet, the resultant in these cases would be a mere epiphenomenon since, although the resultant is ontologically different from its causal components, it is not causally different from them insofar as it tends towards the same manifestation as they do. Mumford and Anjum describe these cases as if the resultant and the component were "somehow the same thing under different guises" (Mumford and Anjum 2011, 42).

Unlike the previous cases, which are a form of actual overdetermination insofar as both the components and the resultant tend towards the same manifestation, Mumford and Anjum recognize another kind of cases that could be wrongly considered to be cases of overdetermination. These are cases in which it *seems* that both the causal components and the resultant tend towards the same manifestation, but what it is really occurring is a case of 'temporally overlapping causal processes':

> During a single simultaneous causal process, therefore, a number of other causal powers will be instantiated, though some only fleetingly. A new and distinct causal process could be triggered, therefore, if one such power meets its mutual manifestation partner(s). What we will have, therefore, is a series of temporally overlapping causal processes, each of which starts when an arising power -one that is instantiated in some other process- meets a new partner (Mumford and Anjum 2011, 126).

Temporally overlapping causal processes occur when there is a causal process (CP_2) that is enabled by the powers that result from a causal process (CP_1), such that its effect is manifested simultaneously with CP_1 occurring. Due to the simultaneity of the processes involved, this kind of situations could be conceived as if the powers of CP_1 were causing the resultant of the causal process CP_2. However, this is an incorrect way of interpreting this situation. According to Mumford and Anjum, what is really occurring in this case is that the resultant of the causal process (CP_1), R_1, has already met its triggering conditions (a.k.a its mutual manifestation partner(s)), such that it has already been triggered giving rise simultaneously to a different causal process (CP_2) that has its own effect, R_2. CP_1 and CP_2, therefore, are two different causal processes with different effects. The resultant of CP_1, R_1, insofar as it is triggered, starts a new causal process CP_2, that has a different resultant, R_2, and that is simultaneous with CP_1. In temporally overlapping causal processes, therefore, what happens is that one causal process offers or provides the powers that are necessary for another causal process to occur (Mumford and Anjum 2011, 127; Anjum and Mumford 2018).

4. A causal dispositional account of fitness

As stated above (§ 2.2), one of the main problems regarding PIF concerns its difficulty to properly distinguish what fitness ontologically is (ecological fitness) from how it is measured (predictive fitness) (Matthen and Ariew 2002). Although for PIF there is a difference between both concepts, fitness ends up equated to predictive fitness, insofar as the expected number of offspring that an organism would leave in a particular environment and population predominates over the ecological characterization of fitness. In this section, I propose an ontological characterization of ecological fitness developed in the metaphysical framework of DTC previously introduced. Following this framework, I propose a definition of fitness that summarizes what I term 'the causal dispositional account of fitness' (hereafter CDAF):

> CDAF_{df}=Fitness is considered as a causal disposition resulting from the non-linear combination of environmentally relative functional dispositions oriented towards an effect (surviving and reproducing) which is reached once the combination of these dispositions exceeds a certain threshold (see Fig. 5).

As I will argue, CDAF has several advantages over PIF and the supervenient account of fitness. Firstly, it accounts for the distinction between ecological and predictive fitness; secondly, the complex and sensitive character of fitness as a disposition that is relative to a particular environment and population is highlighted; finally, the dynamical character of fitness as a property of organisms throughout their life-cycles is also addressed.

In the remainder of this chapter, I will explain in detail each of the conceptual elements involved in this definition and will develop its conceptual potential. Firstly, I will present the complex and dynamical character of fitness as a disposition of organisms (§ 4.1). Secondly, I will highlight the relational character of fitness with respect to the particular environment and population in which the organism is embedded (§ 4.2). Finally, I will clarify the distinction between fitness, expected fitness, and realized fitness within the framework of CDAF (§ 4.3).

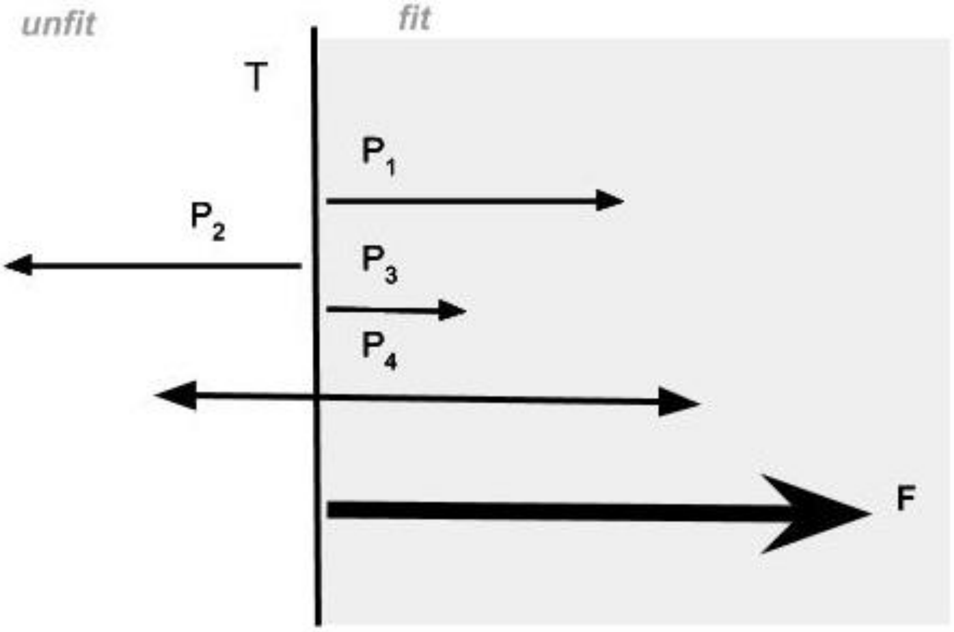

Figure 5. Fitness as a causal dispositional property. "The P_n vectors refer to those dispositional properties whose combination makes an organism fit. Following the DTC, vectors have two properties. First, they have different intensities, as indicated by the different lengths of the arrows. Second, vectors have different directions, indicated by the direction into which the arrowhead is pointing. Functional dispositions are oriented towards (P_1, P_3, P_4) the disposition to survive and reproduce. P_2 is a countervailing factor, that is, a kind of malfunction or trait whose effect is not disposed towards but away from fitness. P_4 reflects a tension in the orientation of that disposition. A sexual trait, for instance, can be oriented towards reproduction (arrow to the right) but be detrimental to the survival of the organism (arrow to the left). Fitness (F) is the resultant of the combination of these dispositions, and is manifested once it has exceeded a certain threshold (T)". (Figure and figure legend reproduced from Triviño and Nuño de la Rosa 2016, 4).

4.1 Fitness as a complex dynamical disposition

In my view, one of the main weaknesses of PIF is that fitness is defined as a single disposition that is reduced to different physical traits of organisms. Conversely, I understand fitness as a complex disposition that results from the combination of other dispositions, namely the functional dispositions that characterize an organism (e.g. the heart pumping blood, the liver producing bile, the legs running…). These functional dispositions, therefore, are the bases of an organism's fitness. Concerning the functional dispositions that characterize an organism, it has been argued that they can be, in fact, reduced to physico-chemical properties such that these 'lower-level' properties ultimately are the real bases of fitness. This is precisely the position defended by Rosenberg's supervenient account of fitness and Mills and Beatty's PIF. As previously seen (§ 2.1 and § 2.2), these accounts are in connection to reductionist approaches to biology that defend the idea that "the functions of a living system are explained solely in terms of the chemical properties of its parts, and these chemical properties will, in turn, be explained by the physical properties of even smaller parts" (Rosenberg 1983).

In contrast, I follow those authors that claim that reducing the function of an organ or a trait to its physico-chemical properties implies losing the function itself. In other words, insofar as we merely appeal to physico-chemical properties, we are not able to know the function that the organ or trait perform (Kauffman and Clayton 2006). Furthermore, due to the complexity of living beings, it is not clear that this kind of reduction could be possible. As John Dupré and Daniel Nicholson state:

> A necessary condition for reductionism is that it must be possible to treat entities independently from one another and to consider

> their structure and constitution independently of the context in which they exist (Dupré and Nicholson 2018, 27).

Yet, this is not the case concerning biological functions. In order to know the function of an organ or trait, it is necessary to take into account both the role it plays within the organism and the functions of other traits and organs that might affect how its own function is performed. Therefore, we can conclude that, "in at least this limited sense, biological explanations necessarily move beyond the explanations at the physical-level alone" (Kauffman and Clayton 2006, 504).

There is a wide debate in philosophy of biology concerning the characterization of biological functions. The problem for conceptualizing biological functions lies in the difficulties for addressing both the explanatory and the normative role of biological functions. In this regard, it is generally considered that an account of biological functions should be able to explain the current presence of the function bearer, as well as to identify what effect a trait is supposed to produce such that cases of malfunctions and accidents could be recognized. Historically, the debate has been dominated by two main positions: etiological and dispositional accounts.[23]

Etiological accounts, or 'selected effect theories' (see Wright 1973; Millikan 1984, 1989; Godfrey-Smith 1994), are characterized for explaining the function of a trait in terms of its causal history. In this regard, it is the effect that the trait performed in the past that has caused its current presence in the organism. Generally, etiological accounts appeal to natural selection as the causal process that explains the current presence of the trait (Millikan 1984,

---

[23] Here I will only briefly present some of the theories regarding biological functions that have been offered. For a general introduction to this debate and to the different accounts that have been offered, see, for instance, Mossio et al. 2009, or Saborido 2013.

1989). The trait was selected due precisely to the effect it performed. Etiological theories of functions, such as the selected effect theory, are able to account for both the normative as well as the explanatory role of functions (Mossio et al. 2009; Saborido 2014). Concerning the normative role, the function of a trait refers to the effect that was selected in the past due to the contribution it made to its bearer. This is what Ruth Millikan calls the 'proper function' of a trait (1989). Within this framework, therefore, functions could be distinguished from malfunctions and accidents: a malfunction occurs when the trait does not accurately perform the effect (function) it was selected for, whereas accidents refer to those effects that a trait performs besides its selected effect or function. Regarding the explanatory role, etiological accounts of dispositions are also able to explain the current presence of a trait in an organism: the trait is currently present in the organism since it was selected through the function it performed in the past and that contributed to its bearer.

Etiological accounts, notwithstanding, have been confronted with different problems (Davies 1994; Cummins 2002). The main one refers to what has been called the 'epiphenomenalism' of etiological functions (Christensen and Bickhard 2002). According to this criticism, the function that etiological accounts attribute to a trait is independent of the current role the trait plays for the system in which it is embedded (see also Gould and Vrba 1982). Godfrey-Smith (1994) has also developed this idea when claiming that, insofar as a function refers to the past selected effect of a trait, it is possible for that trait to have a function without currently performing it, such as vestigial traits.

Dispositional accounts of functions, notwithstanding, do not have the problems attributed to etiological accounts insofar as they characterize functions as the effects of a trait that causally contribute to a distinctive capacity of the system it belongs to (Mossio et al. 2009). An example of a dispositional account is the systemic approach (Cummins 1975), where "a function *F* is the contribution of a process *P* to a distinctive higher-level capacity *C* of the system *S* to which it belongs" (Moreno and Mossio 2015, 66). There are different forms of conceiving the higher-level systemic capacity. Some authors, for instance, have identified it with survival and reproduction (Boorse 2002). Thus, pumping blood is the function of the heart since, in this way, it contributes to the survival of the organism by supplying nutrients and removing wastes in its circularity system. John Bigelow and Robert Pargetter have similarly argued that "something has a (biological) function just when it confers a survival-enhancing propensity on the creature that possesses it" (1987, 108).

Dispositional accounts of functions are able to associate a function with the current activity a trait performs for a higher-level capacity of the system. However, they are not able to address the normative role of a function. Since a function is whatever contribution a trait makes to a higher-level capacity of the system, it is not possible to talk about 'proper functions' of traits, and thus, dispositional accounts cannot properly distinguish among functions, malfunctions, and accidents (Mossio et al. 2009).

Due to the deficiencies of both etiological and dispositional approaches, some authors have proposed a pluralist account to biological functions (Millikan 1989; Godfrey-Smith 1993, 1994) where both approaches need to be considered in order to explain different aspects of biological phenomena. Alternatively, unitary accounts argue for the possibility of elaborating a theoretical framework that covers the advantages of both etiological and dispositional accounts of biological functions (Kitcher 1993; Walsh 1996; Mossio et al. 2009).

It is not the purpose of this thesis to discuss the advantages and disadvantages of the different accounts of biological functions that have been proposed. Yet, CDAF is akin to the dispositional approach to biological functions. In this sense, a phenotypic trait has a function insofar as it contributes to the capacity of the organism to survive and reproduce, that is, to its fitness. These functional dispositions constitute the causal basis of an organism's fitness. Besides functional dispositions, there are also other traits of organisms that might affect their fitness, insofar as they have effects that dispose away from fitness (see Fig. 5, P_2). These effects refer to what Mumford and Anjum call 'countervailing factors' (recall § 3.2). Insofar as they do not dispose toward the systemic capacity of fitness, these effects cannot properly be called 'functions'. Yet, since these traits might also affect how fitness results, that is, how fit an organism is regarding the environment and population in which it is embedded, they need to be taken into account besides the functional dispositions that constitute the causes of an organism's fitness. I will shortly say more on this point. Now, I will illustrate how the features that characterize causality in DTC, such as simultaneity, threshold, intensity, and direction are given in CDAF and help to shed light on this biological notion.

Simultaneity

As previously stated (§ 3.2), in DTC causes and effects are simultaneous in the sense that once a disposition is triggered some effect is produced. Simultaneity is also given in the case of fitness, where the different dispositional components that give rise to the fitness of an organism are also simultaneous with fitness itself. Thus, a healthy organism, whose heart pumps, or whose legs run when its nose smells a predator, is simultaneously a fit organism. I will dwell more on this simultaneous character of fitness and its basis in Chapter 3, when exploring fitness as an emergent disposition.

Threshold

The notion of threshold, as conceived of in DTC to illustrate that a given effect has been produced and that there is not a necessary connection between causes and full effects, plays a crucial role in CDAF. In particular, due to its compositional character, different thresholds might be recognized in the process that gives rise to fitness. On the one hand, the different functions that characterize an organism have different thresholds that are associated with their corresponding goals. Once the threshold is reached, the particular effect associated with the organismal function is manifested. Limbs, for instance, are disposed to run and this disposition is manifested once the muscles and bones that make up a limb have reached a certain developmental stage and encounter an environmental circumstance that leads them to run. On the other hand, organismal functions do not act in isolation. Their particular goals act combined with the other functional dispositions that also compose fitness. Thus, the particular goals of organismal functions are subordinated to the systemic goals of the organism, namely to survive and reproduce. For CDAF, therefore, the fitness of an organism is only reached when the *combination* of the functional dispositions oriented towards particular systemic capacities achieves the *fitness threshold*. In this view, fitness is thus the resultant, not of a single functional disposition acting in isolation, but of the combination of the different functional dispositions acting together. Hence, running is a disposition that, combined with other dispositions, such as pumping blood or seeing, is oriented towards the systemic disposition of fitness, insofar as running contributes to the survival and reproduction of the organism.

Finally, there are also different thresholds that can be recognized regarding fitness as a resultant. As noted above, once the combination of functional dispositions reaches a certain threshold, the resultant fitness is manifested. This resultant is a new disposition that has its own effects associated with it, namely the survival and reproduction of the organism. In this sense, two thresholds need to be distinguished when talking about the disposition of fitness: the *survival threshold*, which is reached whenever there is a proper combination of biological functions that allows the organism to survive (an alive organism is simultaneously a fit organism), and the *reproductive threshold*[24], which presupposes the first threshold and is reached once the organism lives to maturity and reproduces.

The recognition of the two effects of fitness, i.e. survival and reproduction, is an important feature of CDAF as compared to PIF. As stated above (§ 2.2), although PIF characterizes fitness as a disposition to survive and reproduce, it ends up only considering the expected number of offspring of an organism when calculating levels of fitness. PIF defenders might insist on their view by pointing out that the survival of the organism is implicitly considered insofar as, in order to reproduce, the organism needs to be alive (Sober 1993). While I agree that the reproductive effect presupposes the survival effect, I think that it is crucial to keep these two effects of fitness differentiated. The reason is that although there are species of organisms for which reproduction is the relevant aspect to consider, there are also other

---

[24] I will focus in more detail in the reproductive threshold when distinguishing between the full effect of fitness as a disposition to reproduce and other effects that are also associated with it (§ 4.3).

species, e.g., some clonal and colonial organisms, symbiotic communities..., for which survival is more important than reproduction (Bouchard 2008, 2011). Classical evolutionary explanations, for instance, have contributed to increase the importance of reproduction, insofar as they are mainly based on the study of sexually reproducing organisms and the populations they compose.[25] However, not all kinds of biological entities do sexually reproduce, and therefore, they do not form populations in the sense population geneticists consider, i.e. Mendelian populations.

In some cases, such as those concerning species that reproduce asexually, evolution is not accurately understood in terms of reproductive success (Bouchard 2011; see also Ariew and Lewontin 2004). In these cases, adaptive change is produced through the slow accumulation of changes among the parts of the system, rather than through intergenerational change within populations. In clonal species, for instance, there is a problem regarding how should 'individuals' be conceived in them. This question concerning individuality (a problem widely discussed in philosophy of biology (see Wilson 1999; Clarke 2010, 2013; Pradeu 2016)) puts in jeopardy the measurement of fitness in terms of reproductive success. In asexual species, it is not really clear whether there are different individuals produced or whether there is only one individual that is *growing* by creating more parts of itself. Ariew and Richard Lewontin introduce the problem in the following terms:

> The problem that has plagued evolutionists who deal with organisms that have both sexual and vegetative reproduction is

---

[25] There are different forms of understanding the concept 'populations' in contemporary philosophy of biology. It is not the purpose of the thesis to enter into those debates. However, see Sober 1993; Brandon 1990; Walsh, Lewens and Ariew 2002; Matthen and Ariew 2002; Millstein 2006, 2009; Walsh 2007 or Godfrey-Smith 2009, as examples of different forms of conceiving populations.

> how to count ramets and genets in assigning reproductive fitness. Do all the ramets of a single original stem count as belonging to a single individual or is each to be counted as a separate individual? It might be argued that since the ramets are all connected as a single body, they are collectively one individual. But is the occurence of a break in the underground stem sufficient to produce a new individual for accounting purposes? Moreover, the problem exists for trees. A tree consists of a large number of flowering stems connected together by branches and a trunk. Why should it matter that these flowering stems are connected above ground rather than below? (Ariew and Lewontin 2004, 360).

Bouchard is aware of this problem. Yet, he recognizes that, although there might be cases of asexual species in which reproduction is taking place, this is not the case in all of them, and our account of fitness should be able to address this kind of cases as well. In order to illustrate this idea concerning cases of asexual species that do not properly reproduce but grow, Bouchard appeals to an example (Bouchard 2011, 112): the Quaking aspen (*Populus tremuloides*), a species of plant that can reproduce asexually by sending runners below ground that grow. Instead of reproducing, these trees survive during thousands of years by growing. The different branches that grow seem to be genuine trees, such that a clonal grove look like a forest of individual trees although there is only one unitary tree. In cases of asexual reproduction such as this one, selection acts on the parts of a growing individual, not on individuals within a population. This is so since, during the clonal phase, the organism produces more parts of itself, among which there is differential persistence, i.e., survival. In Bouchard’s terms:

> I argue that there is variation in the case of the clone, but not variation among individuals in a population. The variation is amongst the parts of the individual and there is selection amongst the parts of the individual (there is competition between ramets for soil nutrients, i.e. intra-individual competition) (Bouchard 2011, 112).

I agree with Bouchard's consideration that an account of fitness should be able to address this kind of cases. CDAF fulfills this requirement insofar as survival is not implicitly considered in relation to reproduction, but both survival and reproduction are recognized as different effects of fitness.

Direction and Intensity

The metaphysical framework of Mumford and Anjum's DTC also allows making sense and disentangling the complexity of fitness by means of the different directions and intensities by which the particular components of fitness are characterized.

With regard to the *direction* of the functional dispositions, I have endorsed a dispositional notion of fitness according to which functional dispositions are combined and oriented towards fitness. Nonetheless, as previously stated, not every effect of a trait of an organism necessarily contributes in a positive way to its fitness. Those effects that dispose away from fitness are not properly the causes of fitness. Yet, they are also part of the cause, insofar as they are countervailing factors that might affect how fitness is given (§ 3.2). Countervailing factors might refer to malfunctions such as those resulting from the disrupting effects of a malformed or a sick organ. Well-formed and healthy traits may also prove to be maladaptive in a given environment. A classic example regarding peppered moths could serve to illustrate this idea. Thus, the black color of a peppered moth can be adaptive in those environments where the colors of the tree trunks are black, such that the peppered moth camouflages and birds are not able to detect and therefore, to eat it. However, in those environments where the colors of the tree trunks are not black but covered in lichen, for instance, black peppered moths would be more easily spotted out and eaten.

Besides these examples of countervailing factors, I claim that, in the case of fitness, and due to its complex character as a biological property, there can be traits that have both a functional disposition that contributes towards the fitness threshold and an effect that disposes away from it. Sexual traits are an example of this. They can prove to be a hindrance to an organism insofar as they might be oriented towards the goal of reproduction but also against the goal of surviving (see Fig. 5, P_4). A well-known example in evolutionary biology refers to the long tail of many male birds. In some species of birds, to have a long tail is useful to attract females and mate. Thus, having a long tail contributes to the goal of reproduction. Yet, having a long tail also enhances the difficulties for the male bird to fly (both high and fast), and thus, it is easier for him to be caught by a predator. Having a long tail, therefore, is detrimental for the goal of surviving.

Concerning sexual traits, some authors have argued that they are not true adaptations. Sexual traits have been favored by sexual selection insofar as females prefer them, but they do not perform any real function to the organism that bears them (Andersson 1982; Futuyama 1986). Conversely, I agree with those authors who acknowledge that sexual traits are adaptations insofar as they have a function (Godfrey-Smith 1994). In this case, sexual traits contribute to the fitness of an organism since it is a disposition for both survival and *reproduction* (recall here the above idea concerning the two effects of fitness). In the example, therefore, the function of the long tail is to attract females, and this allows to increase the fitness of a particular organism by increasing its capacity to reproduce.

Concerning *intensity*, the functional dispositions (as well as the countervailing factors) that characterize an organism might vary in this regard as well, which implies that the intensity of an organism's fitness might also vary. Since fit organisms are embedded in an environment and population, there are environmental and populational factors that can affect their functional dispositions by altering the intensity with which they are manifested (I will say more on this in Chapter 3). Nonetheless, despite the different intensities by which a functional disposition might be manifested, there are 'viability constraints' (Barandiaran and Egbert 2014) that define the limits of the intensity that the functional dispositions that characterize an organism might tolerate, such that the survival threshold of the combination of functional dispositions can be reached and the organism can keep alive. Each dispositional function has its own viability constraints (for instance, the human heart might fail or not work properly when beating at more than 220 beats per minute) and they vary depending on the organismal and environmental context. Functional dispositions, furthermore, are not only affected by environmental and populational factors, but they also vary in intensity depending on the particular stage of the life cycle of the organism. Thus, the disposition to produce eggs is not manifested with the same intensity during the whole life of a female, for instance.

Dynamism

Changes regarding the functional dispositions that characterize an organism and the intensity how they are manifested make fitness to be a dynamical property of organisms that changes at different stages across their lifetime. In fact, one of the difficulties for characterizing biological dispositions derives from the dynamical character of the biological organization. The morphological, physiological, and behavioral properties of an organism are not static but change along the life history of the organism (see Fig. 6). The functional dispositions that make an embryo fit are not the same as those that make an adult fit. Some traits are essential for the survival of an embryo (e.g. the chorion, or the fetal part of the placenta in mammals), whereas others that are not present at the embryo stage (e.g. some sexual traits) will become vital for the reproduction of an adult. The resultant fitness evolves throughout the development of the organism, and, in this sense, fitness is not only an "instantaneous property", but also "a property of the organism's life history" (Brandon 2005, 166).

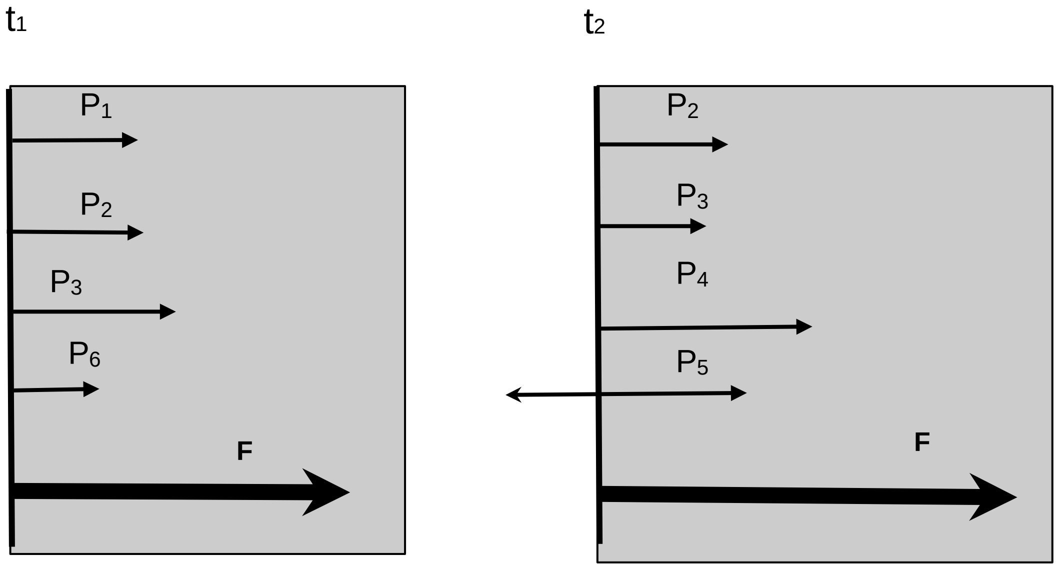

Figure 6. The development of fitness. "Each vector diagram represents a causal situation at a particular time (two different developmental stages, t_1 and t_2) and includes those powers that are causally relevant for the fitness of the same organism. Figure t_1 represents fitness in an embryo. Figure t_2 represents fitness in an adult. The functional dispositions that give rise to fitness change along the life history of an organism: P_1 and P_6 are traits that embryos have but not adults, while P_4 and P_5 are traits (e.g. sexual traits) that only mature organisms have" (Figure and figure legend reproduced from Triviño and Nuño de la Rosa 2016, 8).

4.2. The context-sensitivity of fitness

The role that extrinsic factors can play concerning the fitness of an organism is one of the main debates around PIF (§ 2.2). According to Sober, for instance, population size is an extrinsic factor that need to be considered in the calculation of fitness values. Yet, this consideration challenges the characterization of fitness as a disposition: since calculating the degree of change in fitness in future generations needs to consider the population size, predictive fitness is not based on individual, intrinsic propensities of organisms, but needs to include extrinsic population factors as well (Sober 2001). In the same line, Millstein has argued that propensities are not inherent to individuals, but they are always relative to the environment and population (Millstein 2016). Propensities are always relational properties within a particular context. In this regard, I consider that, in the same way as biological functions can only be properly defined and characterized within the wider context of the whole organism they belong to, the particular fitness of an organism has to be considered in the wide framework of the particular population and environment it is embedded in. The disposition of organisms to survive and reproduce, therefore, is relative to the environment and population

Regarding the environment, an organism's fitness is affected by the factors of the environment wherein it is embedded. Environmental factors, such as food supply, the number of predators, or the weather conditions, might affect the survival and reproductive capacities of the organism. They all influence the way fitness, as a disposition to survive and reproduce, is manifested. Environmental conditions do also affect ecological fitness by affecting the functional dispositions that are the bases of individual fitness.

For instance, a light strike might damage the lungs of a bird such that it cannot properly breath, thus modifying the way the whole fitness is manifested as well.[26] Furthermore, it is not only that functional dispositions vary in relation to the environment, but environments wherein organisms are embedded are also altered by the influence of organisms themselves, insofar as they are niche constructors, i.e. they modify their selective environments (Odling-Smee et al. 2003). Some examples of niche construction are "the building of nests, burrows, mounds, and other artifacts by animals; the alteration of physical and chemical conditions; the creation of shade, influencing wind speed; and the alteration of nutrient cycling by plants" (Laland et al. 2016, 192). In these cases, the changes in the environment caused by the intervention of the organism might also produce new changes in its functional dispositions, such that there is a reciprocal influence between the functional dispositions of organisms and the environments wherein they are embedded.

Regarding populations, the size of the population can also affect the survival of the organism, for example, in those cases wherein food resources are scarce and it is necessary to struggle for them. Furthermore, population size also affects both the possibilities for an organism to find a mate to reproduce, and its reproduction. In the latest case, other organisms in a population can affect how fitness is manifested in reproduction. In the same way as several dispositions need to work together to produce a particular effect, fitness can only give rise to its reproductive effect in combination with the fitness of other organisms. There are many ways in which other organisms might affect each other's reproduction. In particular, in sexually reproducing species, the fitness of an organism depends on the availability, fertility, and

---

[26] Although, as I will make explicit in Chapter 3, this might not be so due to developmental plasticity.

compatibility of individuals in the breeding group of the organism. In this case, fitness is not a property that works in isolation since, in order to be manifested in a certain number of offspring, a fit organism needs to sexually interact with another organism (Millstein 2016).

4.3. Fitness, expected fitness and realized fitness

In presenting the actualist account of fitness, I have shown that one of the problems that philosophers of biology have attributed to it refers to the so-called mismatch problem (§ 1), i.e. the fact that biologists do not seem to use the concept of fitness in the same way as they define it. Organisms that have a higher number of offspring in a given environment and population are not always the fittest ones. PIF attempts to solve this problem by considering fitness as a disposition. Yet, insofar as in PIF fitness is reduced to predictive fitness, there is not an explanation of how ecological fitness relates to both expected and realized fitness. In this section, I explain in detail how the proposed CDAF accounts for the differences between fitness, expected fitness and realized fitness.

4.3.1. Fitness and realized fitness

For CDAF, realized fitness is the effect or the manifestation of the fitness disposition, namely the survival of the organism as well as its number of offspring in a given environment and population. These effects occur when fitness is exercised and reaches its own threshold. The survival threshold, as we have seen, coincides with the fitness threshold insofar as a fit organism is simultaneously a surviving organism. The reproductive threshold, notwithstanding, is reached once the organism actually reproduces.

However, due precisely to its context-sensitivity, fitness might fail to reach the threshold that allows an organism to reproduce. In this kind of cases, organisms do not leave offspring. If these cases are possible, then how can fitness have a causal power? This puzzle can be solved within the framework of DTC. According to DTC, even in those cases in which the disposition fails to reach its threshold, some causation has still occurred insofar as the disposition has been triggered (Mumford and Anjum 2011, 74. See § 3.2). Thus, from the first moment that a particular disposition is triggered, some effect is being produced too, although it might not be the full effect associated with it.

In the case of fitness, even when an organism fails to reproduce, it is still contributing to the function of surviving and reproducing throughout its lifetime. Biological functions do not only result in the survival of the individual organism, but some of them also produce effects in the organism which contribute to its reproduction. In CDAF, when the cells of a blastula differentiate between somatic and sex cells, or when an organism reaches a fertile stage, the organism is moving towards the effect of reproducing. In this sense, we can claim that fitness is disposing towards (and in some sense causing) reproduction even if the full effect associated with reproduction (i.e. having offspring) is not manifested. The effect of fitness as a disposition to reproduce, therefore, is not only manifested in a mature organism, although it is only when the organism is in its mature stage that the disposition to reproduce can reach the (reproductive) threshold associated with its full effect.

4.3.2. Expected fitness and realized fitness

In this section, I argue that my proposed CDAF can offer a more complete account of the distinction between realized and expected fitness. In this frame, expected fitness can be interpreted as the manifestation of fitness in an ideal world, a world which would include the environmentally regular conditions of an organism. However, since in the real world there are always factors which can modify or prevent the manifestation of such disposition, the actual number of offspring of an organism may not coincide with its expected fitness.

In those views of causation in which there is a necessary connection between cause and effect, the existence of exceptions is a problem to the theory, insofar as every time a particular cause is given, the effect associated with it must be given too (Mumford and Anjum 2011; see also § 3.2). In my view, Rosenberg's distinction between fitness and number of offspring (Rosenberg 1978) illustrates the problematic consequences of such an approach to causation. As previously explained (§ 2.1), in Rosenberg's account, the physiological, anatomical, behavioral, and environmentally relative properties on which the level of fitness supervenes, are the causes of the number of offspring of an organism. Rosenberg argues that, since the level of fitness is not the same as the rate of reproduction, two identical organisms can have the same level of fitness without leaving the same number of offspring. However, if there is a causal connection between the properties of an organism on which the level of fitness supervenes and its number of offspring (Rosenberg 1978, 373; § 2.1), it is hard to understand how two identical organisms with the same level of fitness can leave a different number of offspring.

Conversely, CDAF follows DTC in not being committed to a relation of necessity between causes and effects. This is so since there might be interfering factors that might play a role in modifying or preventing the production of the effect. As previously stated, if prevented, the effect is not manifested, i.e. it does not occur. When interfered, the effect is still given but is different from the effect that would have occurred without the interfering factor. In my view, the consideration of interfering factors allows to explain the differences between expected and realized fitness (see Fig. 7).

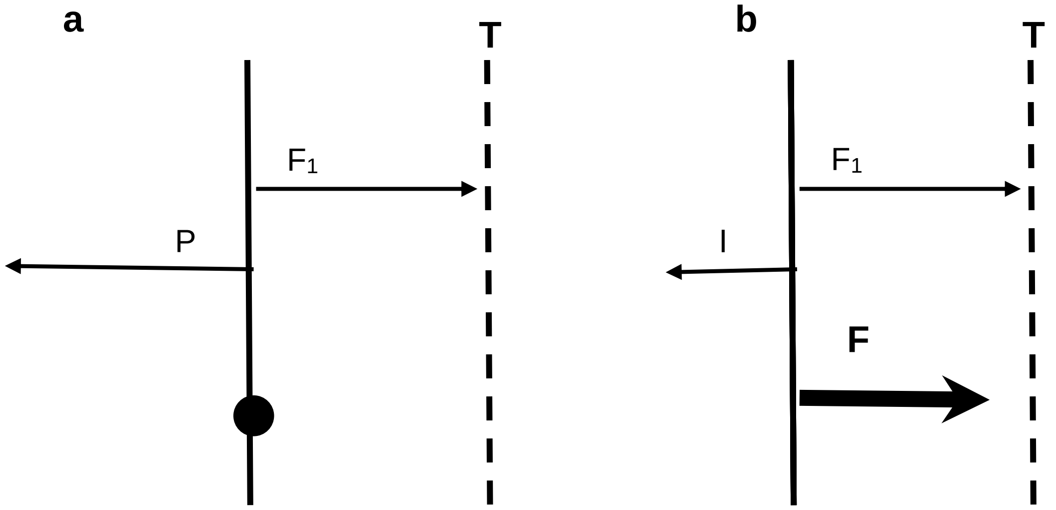

Figure 7. Interfering factors explain the differences between expected and realized fitness. “**Figure 7a** represents the case where a preventing factor (P) prevents fitness to be manifested (black bullet). **Figure 7b** represents the case where an interfering factor (I) affects the fitness of the organism modifying the resultant fitness (F), which is unable to reach the reproductive threshold (T).” (Figure and figure legend reproduced from Triviño and Nuño de la Rosa 2016, 14).

The differences between interfering and preventing factors can be illustrated by means of the example of twins (see Fig. 7a). When both twins are standing together in a forest, it is possible to claim that both have the same expected fitness and therefore, the same level of fitness, insofar as both are identical: both twins are in the same environment and population and have the same functional dispositions, whose combination has reached the threshold that gives them the same disposition to survive and reproduce. However, although the expected fitness is the same, the realized fitness is different, since there is an interfering factor (the lightning strike) that affects one of the organisms but not the other.

In this case, the lightning is an interfering factor that prevents the manifestation of fitness by killing the organism. By means of the lightning, every functional disposition that characterize the organism has been affected such that the fitness threshold cannot be reached (Fig. 7a). At the same time, the preventing factor also alter the intensity by which the fitness of the surviving twin is manifested. Due to the contextual sensitivity of fitness, the surviving twin now has one competitor less, which might increase its survival and reproduction.

As stated above, interfering factors might affect the manifestation of a disposition not only by preventing the effect but also by interfering with it, that is, by changing the form how the effect is produced. In the example of the twins, this idea can be illustrated by considering that the unlucky twin does not die but it is seriously damaged due to the lightning strike. In this case, different lesions in certain (encephalic, cardiac, pulmonary…) functional dispositions that characterize the organism might affect the way its fitness is manifested by making it to have a less number of offspring (or to live less years) than the one it would have had without the interfering factor occurring. In the case of the interfering factor that interferes with fitness without preventing it, therefore, a particular effect of fitness is still manifested (realized fitness), although this effect differs from the one that had been previously calculated without considering the lightning event (expected fitness) (Fig. 7b).

Recall that in PIF, Mills and Beatty also acknowledge the possibility of fitness being affected by some ‘disturbing factors’ (Mills and Beatty 1979, 271-272). However, they only think of these factors as preventing, and not also interfering with, the manifestation of fitness. Mills and Beatty, therefore, only consider those factors that avoid the possibility for an organism to leave offspring by killing it, for instance. Furthermore, they especially mention that these factors are not considered in the prediction of the expected fitness of organisms. Only phenotypic traits of organisms are considered when calculating expected fitness. CDAF differs from this approach. On the one hand, CDAF recognizes not only the possibility of fitness to be prevented, but also its possibility of being interfered. On the other hand, interfering and preventing factors might be considered in order to calculate the expected fitness of an organism (although they are not part of the functional dispositions that give rise to fitness), since the organism is given in an environment and population, and these factors might affect its survival and reproduction, as argued above.

In DTC, the possibility of interfering factors explains that some scientific predictions fail or are false (Mumford and Anjum 2011, 137). Predictions are formulated within a particular finite model that *assumes* the presence of some powers while excluding others. However, this model might be inadequate when applied to the real world. A prediction is good as long as it takes into account all the factors that play a causal role in the manifestation of a particular disposition. Yet, since not all the causal elements can be known when making a prediction, some predictions fail. According to DTC, the defeasibility of prediction illustrates that our predictions need to be revised and modified according the new evidence provided. This evidence has to be included in the model in order to improve it and obtain better predictions.

With regard to fitness, the possibility for interfering factors to affect how the fitness of an organism is realized is related to its probabilistic character. The probabilistic nature of fitness (i.e. the fact that the expected fitness predicted by biologists is only the *probable* number of offspring that an organism will leave) might be a consequence of the existence of interfering factors that happen to occur but that were not considered at first by biologists. As previously stated (§ 2.2), two positions have been held regarding the probabilistic character of fitness. The proponents of PIF have generally endorsed an indeterministic interpretation of the fact that organisms have propensities of different strengths to leave various numbers of offspring (Mills and Beatty 1979; Brandon and Carson 1996; Brandon 2005, 154). Conversely, some authors have argued that the probabilistic character of fitness lies on epistemic limitations, and results from our impossibility to know *all* the factors taking part in evolutionary processes. If all the factors that cause the actual level of fitness of an organism (i.e. its anatomical, physiological, and behavioral traits in relation to the environment) were known, the precise prediction of the expected number of offspring of an organism would be possible (Horan 1994; Graves, Horan, Rosenberg 1999; Bouchard and Rosenberg 2004). As a consequence, fitness should not be defined as a probabilistic disposition, since the adduced probabilistic character of fitness does not reflect a real aspect of fitness, but an epistemic limitation

Concerning CDAF, notwithstanding, the interfering factors that were not considered when calculating expected fitness but that play a role in the manifestation of its effect cannot be said to be *always* causally relevant. That is, not all interfering factors should be included in our predictive model once they take place, as suggested by Mumford and Anjum. As Marshall Abrams has argued in detail, not all the elements given in a spatio-temporal region are relevant to calculate expected fitness. Rather, we should only include those environmental and temporal factors which are *constitutive elements* of the environment of a population (Abrams 2006, 2007). In particular, the roles of environmental factors in natural selection "differ depending on the probability of being experienced repeatedly by organisms, and how environmental conditions combine probabilistically to help determine fitness" (Abrams 2007). Thus, returning to the example of the twins, the low probability of being hit by lightning prevents us from including lightning as part of the regular environmental conditions of birds, despite it being the interfering factor that alters our predictive model and produces a deviation between the expected and the realized fitness.

The difficulty concerning the kind of elements that we should include in our predictive model lies, therefore, in determining how many environmental factors are part of the regular environment that influences fitness (Bouchard and Rosenberg 2004; Abrams 2007). However, I think that this should not lead us to abandon the need to specify a 'normal environment' in our causal models of fitness. The stipulation of a normal environment is necessary to distinguish selection from accidental evolution (Walsh 1996). In the case of twins, we can say that the fact that the lightning strike hit one of the twins was a random and unusual event that should not be included in our definition of the normal environment of the population these twins belong to. Therefore, causal models of fitness should not be modified every time a random event creates a situation in which expected fitness differs from realized fitness. Of course, the random or constitutive character of an environmental factor depends on the species. There might be species which are adapted to the regular occurrence of bushfires caused by lightning, for example. In this case, lighting will not be considered as a random and unusual event, but as a part of the regular environmental conditions of that species. In this case, the lighting should be included in our predictive model. Due to the possibility of random and unusual events to occur, i.e. events that are not part of the normal environment of an organism, from the CDAF perspective here defended, this kind of events should not be included in our predictive models of fitness.

5. Conclusions

In this chapter, I have offered a causal dispositional account of fitness wherein fitness is conceived of as *a causal disposition resulting from the non-linear combination of environmentally relative functional dispositions oriented towards an effect (surviving and reproducing) which is reached once the combination of these dispositions exceeds a certain threshold.* I have argued that this account has several advantages over the existing definitions of fitness. In particular, Mumford and Anjum's theoretical framework allows developing the propensity interpretation of fitness in new directions that overcome some of the difficulties associated with the version originally formulated by Mills and Beatty. Firstly, this proposal accounts for some specific traits of living beings (namely complexity, contextuality, and dynamism) that PIF does not deal with. Secondly, my definition of fitness offers a clearer distinction between the concepts of fitness, expected fitness, and realized fitness. In this regard, fitness is conceived of as a causal dispositional property of organisms.

Expected fitness refers to the mathematical formulation of the manifestation of an organism's fitness in an ideal world, that is, a world which include the environmentally regular conditions of an organism. Finally, realized fitness refers to the actual manifestation of an organism's fitness (i.e. its survival and reproduction) in a particular environment and population. The possibility of distinguishing among these concepts solves some of the problems faced by current definitions of fitness.

Chapter 3:
The Emergent Character of Fitness

> *"An organized being is then not a mere machine, for that has merely moving power, but it possesses in itself formative power of a self-propagating kind which it communicates to its materials though they have it not of themselves; it organizes them, in fact, and this cannot be explained by the mere mechanical faculty of motion."*
>
> (Immanuel Kant 1790)

Some philosophers of biology have identified the causal power of fitness with the physical and chemical properties of organisms. Rosenberg, for instance, claims that fitness is "*nothing more than* having a certain combination of anatomical, physiological and environmentally relative properties" (Rosenberg 1978, 374). In the same line, the upholders of PIF consider that every manifestation of fitness can be, in principle, explained and expressed in terms of the different components and subcomponents of the fitness disposition, i.e. the physical and chemical properties of the organism that underlie its viability and fertility levels. However, these accounts acknowledge that the concept of fitness is at least epistemologically needed, insofar as "there seems to be no particular function of components and subcomponents in terms of which fitness can universally be expressed" (Beatty and Finsen 1989, 22). In other words, the concept of fitness is maintained due to its epistemological role in *explaining* the relative success of organisms in producing offspring. As Rosenberg has claimed, without the concept of fitness, evolutionary explanations would be "unmanageable in length, and incapable of predicting biologically interesting consequences of fitness levels" (Rosenberg 1983, 459). The concept of fitness, therefore, has explanatory power and, in this sense, should be preserved as a primitive theoretical term.

In contrast, in the causal dispositional account of fitness I have defended in the previous chapter, fitness does not only have explanatory role, but also causal power. In CDAF, fitness is characterized as the property that makes organisms able to survive and reproduce in a particular environment and population. However, something more needs to be said concerning this causal power of fitness, especially regarding its autonomous or non-autonomous character. In this regard, if the causal power of fitness were, in fact, reducible to the different bases of fitness, i.e. to the causal power of the functional dispositions that characterize organisms, then CDAF would not really differ from the supervenient or PIF accounts of fitness. Conversely, if the causal power of fitness were irreducible, then the ontological status of fitness as a causal autonomous property could be argued for. This is precisely the goal of this chapter.

In arguing for the causal autonomy of fitness, I appeal to the notion of emergence, which is used in metaphysics to conceptualize and make sense of higher-level properties that are causally autonomous. In particular, I will explore whether the features that are generally applied to emergent properties by metaphysicians are met in the case of fitness. In doing so, I will proceed by considering Jessica Wilson's recent review of ontological emergentism (Wilson 2016). I choose Wilson's proposal because my main aim refers to the ontological characterization of fitness, and Wilson explicitly argues that, although there are epistemological characterizations of emergence, all of them are based on two types of *ontological* emergence, namely weak and strong. Wilson proposes a classification of different ways of conceiving the main features of ontological emergent properties, i.e. dependence and autonomy, and illustrates how they match weak and strong ontological emergence. Regarding fitness, I appeal to Wilson's categorization of the different ways of understanding dependence and autonomy and explore which one captures better my view of fitness as a causal disposition.

The structure of the chapter is as follows. Firstly, I introduce Wilson's characterization of weak and strong ontological emergence within the more general framework of emergentism (§ 1). In Section 2, I pay attention to the different ways the notion of dependence is conceived of in emergentism and explore which of them match better the ontological characterization of fitness as a causal disposition. In this regard, I explore whether the relation between fitness and its basis is one of material composition (§ 2.1), modal covariation (§ 2.2), fusion (§ 2.3), non-reductive realization (§ 2.4), or causation (§ 2.5). As I will argue, the kind of dependence that is given between fitness and its basis matches strong ontological emergentism. In Section 3, I focus on the autonomy of fitness. In § 3.1, I introduce some of the ideas concerning the ontological status of organisms and their features from the philosophy of biology. In § 3.2, I offer different arguments for fitness as a causally autonomous property. Finally, I consider a possible objection to this thesis and offer a solution to it (§ 4).

1. Ontological emergence: weak and strong

The idea of emergence has been formulated since ancient times, including Aristotle and Plotinus as precursors of the concept (Clayton 2006). However, its major conceptual development took place during the late 19th and the early 20th centuries after the work of authors such as John Stuart Mill (1843), Henry Lewes (who coined the term 'emergent' in 1875), Samuel Alexander (1920), Lloyd Morgan (1923) or Charlie Dunbar Broad (1925). Although these authors disagreed about how to characterize emergence (O'Connor and Wong 2005; Kim 2006), there were also enough epistemological and metaphysical convergences among them to generate a philosophical movement that has been labeled 'British Emergentism' (McLaughlin 1992, 1997; Kim 2006).

During the late 19th and the early 20th century, reductionist approaches to scientific disciplines such as biology were in jeopardy (Morange 2006), and emergentism appeared as a new way of conceptualizing different natural phenomena which seemed to be irreducible to physical categories. In particular, emergentism was developed as a middle position between reductionism and dualism (in psychology) or vitalism (in biology). In biology, reductionist positions considered that properties such as 'life' could be reduced to physics and chemistry (Baetu 2012, 434).

For the vitalists, conversely, 'life' is explained in terms of entelechy, that is, of something immaterial (what Henri Bergson referred to as the *élan vital*) that distinguishes living beings from inanimate objects, and that prevents the reduction of organismic functions to chemistry and physics (for a historical review on vitalism, see Nicholson and Gawne 2015). Vitalism was criticized by reductionists, who claimed that, on the one hand, it cannot explain why entelechies *operate differently* among different living beings. The entelechy of a salamander, for instance, can restore its amputated arm whereas the entelechy of a man cannot do so (Lillie 1914, 843). And, on the other hand, given that entelechies are *immaterial*, that is, they do not occupy space, it cannot be possible to isolate them through experiments nor to elaborate accurate arguments for their existence (Lillie 1914; Beckermann

1992; Mayr 2004; Nicholson and Gawne 2015). Against vitalism, emergentists argued for a unitary view of the world wherein there are only material entities (Stephan 2002). Against reductionism, they claimed that the material world is not a simple, linear domain, but it consists of a hierarchy of layers that complexly organize its material elements (McLaughlin 1992, 1997).

The British Emergentists established and developed the core idea of emergentism: when physical systems reach a certain degree of complexity, they can manifest new higher-level domain properties that are not possessed by the lower-level components of the systems nor are reducible to them (Kim 2006, 548). The work of C. D. Broad, *The mind and its place in Nature* (1925), is considered the last major work within the British Emergentist Tradition (McLaughlin 1997). Due to the new explanations in physics accounting for some properties of chemical and biological systems so far considered to be irreducible, as well as to new research in complex non-linear systems (e.g. turbulent fluids), emergentism disappeared from the mainstream philosophy between the 1930s and 1970s (McLaughlin 1992; Wilson 2013).

During those decades, emergentism was widely criticized for being a highly questionable, confusing, and incoherent position that could be overcome with the progressive development of science (Pepper 1926; Emmeche et al. 1997), and reductionist positions advocated by the logical empiricists predominated in philosophy of science. This tendency is well illustrated in the thesis of the Unity of Science, according to which all of the so-called special sciences -chemistry, biology, sociology, psychology…- can be reduced, at least *in principle*, to the most fundamental one, i.e. physics (Oppenheim and Putnam 1958; see also Gilbert and Sarkar 2000, 1). According to the reductionist tendency, reduction to physics was conceived of as a constraint for each special science to be properly a science (Fodor 1974, 97).

The attempt of reducing the classical genetic theory to molecular biology is the classical example that serves to illustrate this thesis among philosophers of biology (Oppenheim and Putnam 1958, 22; see also Schaffner 1967, 1969; and Waters 1994).

Since the last decades of the 20th century, nonetheless, the thesis of the unity of science has been widely challenged (Fodor 1974; Maull 1977; Kitcher 1981, 1984b; Mayr 1982, 1988, 2004; Dupré 1993, 2012), and emergentism has reappeared in scientific and philosophical contexts. Within the framework of the philosophy of mind, for instance, emergent positions concerning consciousness started to appear against reductionistic and epiphenomenal approaches. Roger Wolcott Sperry is an example of this tendency:

> Whereas the older interpretations of consciousness as inner aspect, epiphenomenon, or semantic pseudoproblem have remained largely sterile, conceptually and experimentally (e.g., there is no place to go from an epiphenomenon), the emergent interaction scheme is by contrast potentially fruitful (Sperry 1969, 534).

As occurred with the first conceptual development, the philosophical revival of emergentism has its origins in the collapse of reductionism, in this case, of the reductionist positions widely advocated by the logical empiricists (Kim 1989, 1999, 2006; Clayton 2006). Contrary to what was expected, the rejection of the reductionist accounts did not give rise to dualist approaches to reality (Kim 1989), but rather to middle positions between dualism and reductionism, labeled 'non-reductive materialism' (Kim 1989, 32). Accounts, such as those of multiple-realizability, supervenience, or emergence, among others, are examples of forms in which non-reductive materialism has been conceived (Kim 1989; Beckermann 1992).

Currently, the widespread development of the so-called sciences of complexity has also contributed to the increasing defense of emergentism (Goldstein 1999; Goldstein 2004; Witherington 2011). Since its reappearance, metaphysical debates on emergence are oriented to make sense of the basic ideas associated with the concept, namely dependence and autonomy (van Guilick 2001; Clayton 2006; Barnes 2012). As Robert van Gulick puts it:

> (...) the core idea of emergence is that "Xs are more than just Ys" and that "Xs are something over and above Ys". Though the emergent features of a whole or complex are not completely independent of those of its parts since they "emerge from" those parts, the notion of emergence nonetheless implies that in some significant and novel way they go beyond the features of those parts (van Gulick 2001, 16).

Despite its current use, there is not a unified account of what emergence is yet (Kim 2006; Wilson 2016). It is not only that the variety of definitions is useless for clarifying the structure of natural reality (Wilson 2016), but also that the philosophical characterizations of emergence are, in some cases, so complex and abstract that it seems really difficult to conceive of an emergent property in the world (Macdonald and Macdonald 2009). However, in different scientific explanations, especially those regarding complex systems, the concept is widely used, and emergent properties are widely recognized. This fact seems to illustrate that there is an incompatibility between how the concept is defined by metaphysicians and how it is used and conceived by philosophers of science. In philosophy of biology, for instance, the abstract characterization of emergent properties given by metaphysicians is considered to be problematic, insofar as it is unable to make sense of the emergent character of some biological properties, such as the amount of nectar stored in a hive (Mitchell 2012).

The main difficulty concerning the characterization of emergence relies on properly explaining that a higher-level property that depends on lower-level properties of the parts of the system in which it is embedded might be, at the same time, autonomous from those lower-level properties. Different authors have claimed that the autonomy of some emergent properties is a consequence of them having a novel causal power (O'Connor 1994; Kim 2006), that is, a causal power that is qualitatively different from the causal power possessed by the properties of the parts that constitute the system and their interactions. This causal power, furthermore, is said to be downwardly manifested, insofar as the causal influence of emergent properties is exerted on the lower-level properties that constitute the system in which they are instantiated (O'Connor 1994, 97-98).

Since the 1970s, downward causation is recognized as a central component of emergentism. In philosophy of biology, Donald Campbell suggested that the lower level entities of a system can behave in ways that they would not do, due to the constraints created by the higher level organization that they constitute (Campbell 1974). Despite being a central hallmark of emergentism, some authors, in particular, Jaegwon Kim, have argued that it is precisely downward causation that puts emergentism in jeopardy by making it unintelligible (Kim 2006). According to Kim, if emergent properties exert downward causation, then they cannot have an autonomous causal power. Kim argues for this idea by proposing what he calls 'the causal exclusion argument' (Kim 2006, 558). This argument is based on two principles: 1) the *principle of downward causation,* according to which "to cause any property (except those at the very bottom level) to be instantiated, you must cause the basal conditions from which it arises (either as an emergent or as a resultant)" (Kim 1999, 24; see also Kim 1993); and 2) the *causal inheritance principle,* which states that "if a functional property *E* is instantiated on a given occasion in virtue of one of its realizers, *Q*, being instantiated, then the causal powers of this instance of *E* are identical with the causal powers of this instance of *Q*" (Kim 1999, 15).

Based on these two principles, Kim proceeds as follows: let's consider that we have an instantiated emergent property *M*, given at a particular higher level *L*. The instance of *M*, insofar as it is emergent, has a causal power that is downwardly exerted. Thus, in order for this instance of *M* to cause another property *M′* to be instantiated, it needs to cause the instantiation of the basal conditions of *M′*, say *P′*, in the lower level *L-*, from which *M′* can arise (as per the principle of downward causation). However, and due to its emergent character, the instance of *M* also needs the instantiation of basal conditions from which it can arise, *P*. The instance of *M*, therefore, arises from the instantiation of *P* and, due to the causal inheritance principle, *M* inherits all its causal powers from *P*. In this sense, Kim argues, the instantiation of *M* does not have any causal power on its own. And thus, it is the instance of *P* that really causes *P′* to be instantiated. Furthermore, since *P′* refers to the basal conditions of *M′*, it is *P′*, not *M*, that causes *M′* to be instantiated. Kim states this idea as follows: "If an emergent, *M*, emerges from basal condition *P*, why can't *P* displace *M* as a cause of any putative effect of *M*? Why can't *P* do all the work in explaining why any alleged effect of *M* occurred?" (Kim 1993, 208). To claim, for instance, that both the emergent property *M* and its basal conditions *P* are the causes of *M′*, would imply to claim for the overdetermination of causation:

> This appears to make the emergent property *M* otiose and dispensable as a cause of *P′*; it seems that we can explain the occurrence of *P′* simply in terms of *P*, without invoking *M* at all. If *M* is to be retained as a cause of *P′*, a positive answer has to be

> provided. If *M* is somehow retained as a cause, we are forced with the highly implausible consequence that every case of downward causation involves causal overdetermination (since *P* remains a cause of *P′* as well) (Kim 2006, 558).

In order to avoid causal overdetermination, Kim appeals to what he calls 'the causal closure of the physical', according to which, "if a physical event has a cause at *t*, it has a sufficient physical cause" (Kim 2009, 38). In this regard, emergent properties are not considered as causally efficacious, insofar as the lower-level properties on which they depend are the sufficient physical cause. By criticizing emergentism, the pretended upshot of Kim's argument is a defense of reductionism. This is so since epiphenomenalism, which is another option left for the physicalist, "strikes most of us as obviously wrong, if not incoherent" (Kim 2005, 70). The general physicalist result of the exclusion argument, therefore, would be that the reduction of the emergent property to the lower level ones is an inevitable consequence (Morales 2015).

Kim's exclusion argument has been widely criticized, and different alternatives arguing for the causal power of emergent properties have been proposed (see, for instance, Jackson and Pettit 1990; Shapiro and Sober 2007; Macdonald and Macdonald 2009; see also Baysan and Wilson (2017) for some criticisms to these responses). Besides these criticisms, which directly appeal to Kim's metaphysical argument and its conclusions, other authors focus on questioning some of Kim's presuppositions. According to David Witherington (2011), for instance, the problem with Kim's exclusion argument is that the kind of cause that he is considering when referring to 'downward causation' is the efficient one. This kind of cause, which has been the predominant one in science since the 17th century, invokes the antecedent conditions for a phenomenon to explain it. In this regard, since the antecedent conditions of an emergent property are the lower-level properties of the parts of the system in which it is embedded, emergent properties are nothing more than their lower-level realizers, i.e. they do not possess autonomous causal powers. Witherington puts this idea as follows:

> In fact, Kim's charge of causal overdetermination presupposes a causal monism. For Kim, the only real form of explanation is efficient causal explanation, and as a result, the notion of downward causation from Kim's vantage point is framed in the same push-from-behind, antecedent terms as the bottom-up dynamics of microdetermination (Witherington 2011, 72).

As an alternative to this 'causal monism' in which the only relevant cause is the efficient one, Witherington proposes 'causal pluralism', in which efficient, as well as formal, material, and final causes are considered (Witherington 2011). According to Witherington (see also Overton 1991; Tolman 1991; Valsiner 1997), both final and formal causes are the ones that play the main role when talking about the causal powers of emergent properties (I will come back to this idea in § 3). As he claims:

> [...] modern advocates of ontological emergence uniformly reject the notion of strict 'efficient' temporal causality from an independent higher level to a lower one, arguing instead for an identification of downward causation with formal and final cause (Witherington 2011, 75).

Due to the difficulties in providing an accurate definition of emergence that properly makes sense of its creed: dependence and autonomy (Sartenaer 2013), different criteria have been proposed to recognize emergent properties. For some authors, emergent properties are unpredictable, irreducible, and have causal powers (Kim 1999). Others have offered a more extensive list of criteria for a property to be emergent, namely holism, novelty, qualitative difference, not given at lower levels, different laws, and interaction (Humphreys 1997). These criteria, as Paul Humphreys makes explicit (Humphreys 1997), are sufficient but not necessary to talk of emergence: they do not need to be met by all emergent properties, nor do all emergent properties need to share the same features. This implies that different accounts of emergence might be given depending on the criteria one considers emergent properties need to meet (Macdonald and Macdonald 2009).

Despite the extant definitions of emergence, it has been claimed that all of them can be subsumed into two more general categories, namely weak and strong emergence (Chalmers 2006; Clayton 2006). According to David Chalmers, a higher-level phenomenon is weakly emergent with regard to a lower-level domain when the phenomenon is *unexpected* or *unpredictable* given the principles governing the lower-level domain. A phenomenon is emergent in the strong sense when it is not *deductible*, even in principle, from the truths in the lower-level domain (Chalmers 2006). As some authors have argued, the lack of *in principle* deductibility implies that the higher-level phenomenon in question is not reducible to the lower-level domain upon which it depends (Stephan 1999, 2002). Those higher-level properties whose behavior can be deduced from the behavior of the system's parts alone, or from the behavior the system's parts show in simpler systems than the system in question, are reducible. Notice that this does not exclude unpredictable properties to be reduced to their bases if this unpredictability is given *in practice* and not in principle. The fact that a higher level property had not been predicted before its appearance does not exclude the possibility for that property to be reduced to the lower-level properties of the parts of the system in which it is embedded once it appears. In this case, the difficulties for predicting the manifestation of the higher-level property lie on our practical and epistemological limitations. The higher-level property, in this sense, was unexpected or weakly emergent.

Irreducible properties are also unpredictable in principle (Stephan 1999, 2002). According to Achim Stephan, if we cannot reduce a system's property *P* to the properties of the parts that characterize it, or to the behavior and interactions among those parts, then we are not able to predict, from the properties of the parts that characterize the system nor from the relations that exist among those parts, that the system is going to manifest a property *P* (Stephan 1999). Yet, this unpredictability, although given in principle, does not preclude us for the possibility to make other kinds of predictions, since those same systems with same parts and interactions among them will also manifest the property *P*. This is what Kim has called 'inductive prediction' (Kim 1999; Stephan 1999; Sartenaer 2015). Stephan puts this idea in the following terms:

> [...] a property can be unpredictable even though the novel system's micro-structure is predictable. That is the case if the property itself is irreducible: for if systemic properties are irreducible, then they are unpredictable before their first appearance. However, this does not preclude that further occurrences of such properties might be predicted adequately (Stephan 1999, 53).

Due to the general associations between deduction and reduction[27], strong emergence, insofar as it does not allow for even an in principle deduction, is associated with irreducibility, while weak emergence is associated with unpredictability. Weak emergent properties are unexpected

---

[27] For the classical Nagelian account, *reduction* is a relation given between two theories T_1 and T_2 if the laws of the theory T_1 can directly or through the help of bridge laws, *be deduced from* the laws of the theory T_2.

properties, although they can be reduced to the lower-level domain once they appear.

Although there is a tendency to equate weak and strong emergence to epistemological and ontological emergence, respectively (Bedau 1997; Silberstein and McGeever 1999; see also Mumford and Ajum (2011), Chapter 2, § 3.1), I agree with those metaphysicians who argue that weak and strong emergence are *degrees* in which both ontological and epistemological emergence can be given. As van Gulick (2001) and Wilson (2016) have made explicit, epistemological emergence refers to those properties that we, humans, characterize as such due to some limitations we have to know or explain them. They are, in a sense, subjective. Ontological emergence, notwithstanding, refers to a kind of objective properties that are given in the world (van Gulick 2001, 16).

The distinction between epistemological and ontological emergence, therefore, is not accurately captured by equating them to weak and strong emergence, respectively. Furthermore, since in accounts such as Chalmer's, for instance, weak and strong emergence are associated with non-prediction and non-reduction, respectively, both refer, in fact, to two different degrees (namely weak and strong) in which epistemological emergence is given (see Guay and Sartenaer 2016). If an epistemological emergent property is weak, then the property is unpredictable *in practice*. In other words, there is some kind of cognitive limitation or science has not advanced enough to make possible the prediction of the systemic property. However, if the epistemological emergent property is strong, the unpredictability of the systemic property is given *in principle* and not only in practice, which implies that the property is also irreducible.

Generally, unpredictability in principle has been associated with the indeterministic character of some systems. However, it has been acknowledged that this kind of unpredictability is also proper of deterministic systems such as chaotic ones (Sartenaer 2015). Chaotic systems are so sensitive to initial conditions that the predictive algorithms are complex and difficult to manage. In these cases, some authors have argued that the only way to know the future state of the system is by means of simulations (Bedau 1997; Emmeche et al. 1997; Humphreys 2008). Regarding the characterization of epistemological emergence, as we see, epistemological notions such as those of reduction, prediction or deduction play the main role.

Ontological emergent properties are characterized by being, at the same time, dependent on the lower-level properties of the parts of the whole in which they are given, and autonomous from them due to some causal power[28] they have. Regarding ontological emergence, characterizations in terms of causal power, dependence, and autonomy are the relevant ones (Wilson 2016). Since my task in this chapter is to explore whether the causal power that is attributed to fitness when characterized as a causal disposition is autonomous or not, I will mainly pay attention to ontological emergence and its different degrees.

Ontological emergent properties

Ontological emergent properties (henceforth OEP) are mainly conceived of on the basis of causal powers. According to the so-called Alexander's dictum, a

---

[28] As we will see, depending on whether the ontological emergent property is strong or weak, the causal power is understood as novel or as a proper subset of the causal powers of the basis upon which the emergent property depends, respectively.

property is real, i.e. *ontologically autonomous*, when it has causal powers. In this sense, if OEP are real properties in the world, then they must have causal powers. The notion of causal power is, notwithstanding, a problematic one, since it can be differently conceived depending on the ontological commitments one might have regarding properties. Here, I will follow Wilson's account of causal powers, according to which having a causal power means that the bearer of the property has the power, in appropriate circumstances, to enter into causing some effect (Wilson 2002, 2016). Although this account is neutral regarding whether properties are ontologically categorical or dispositional (recall Chapter 2, § 3.1), in this chapter I will follow Mumford and Anjum's characterization of properties as dispositions, i.e. as causal powers (Mumford and Anjum 2011), since it is in terms of dispositions that fitness and its basis are conceived within the causal dispositional account of fitness (CDAF) presented in the previous chapter.

In order for a property to be emergent, nonetheless, it is not sufficient for it to have causal power. This causal power must also be *autonomous* from those of the lower-level properties on which the emergent property depends. Causal autonomy implies ontological autonomy, whereas the reverse does not necessarily follow. A property might have a causal power and therefore, be ontologically autonomous (by Alexander's dictum) without being causally autonomous, insofar as its causal power is also given in the lower-level properties upon which it depends (Kim 2006; Jesper 2006; Macdonald and Macdonald 2009; Wilson 2016). In Kim's words:

> [...] the causal powers the emergentists bring with them must be new and distinctive (remember: emergent properties are supposed to be "novel"), if they were reducible to the causal powers of the base-level properties, they would bring nothing new and would

> have nothing new to contribute to the evolving causal structure of the world (Kim 2006, 557).

In those cases in which there is not causal autonomy, there is epiphenomenalism but not emergence. In fact, the difficult aspect regarding emergence is to properly distinguish it from epiphenomenalism (Collier and Muller 1998). In this regard, Wilson (2016) offers a taxonomy of two different ways emergent properties might be causally autonomous, and therefore, distinguished from epiphenomena. These two ways correspond, in fact, to the two different degrees in which OEP might be given, namely strong and weak (Wilson 2016):

Strong Emergence (SE): "Token higher-level feature *S* is strongly metaphysically emergent from token lower-level feature *P*, on a given occasion, just in case (i) *S* synchronically depends on *P* on that occasion; and (ii) *S has at least one token power not identical with any token power of P on that occasion*" (362. Emphasis added).

Weak Emergence (WE): "Token higher-level feature *S* is weakly metaphysically emergent from token lower-level feature *P* on a given occasion just in case (i) *S* synchronically depends on *P* on that occasion; and (ii) *S has a non-empty proper subset of the token powers had by P, on that occasion"* (362. Emphasis added).[29]

---

[29] Weak emergence refers to the schema associated with non-reductive physicalism (Wilson 2016, 362).

For SE, there is causal autonomy insofar as the emergent property incorporates new causal powers in the world. In cases of WE, conversely, the emergent property does not incorporate new causal powers. Yet, as Wilson claims, the weak emergent property still has a different *causal power profile* since it only possesses a *proper subset* of the causal powers of the lower-level properties it depends on. Insofar as it has causal power, the property is real (as per Alexander's dictum). Furthermore, since the causal power profile is different from that of the lower-level properties upon which it depends, the higher-level property is causally autonomous as well (Wilson 2016, 362).[30]

Concerning the two types of causal autonomy that can characterize OEP, Wilson offers a taxonomy of the different kinds of dependence that can be given between the higher-level property of a system and the lower-level properties of the parts that compose it, namely material composition, modal covariation, fusion, non-reductive realization, and causation. She also illustrates how these kinds of dependence match WE and/or SE. In the following section, I will explain Wilson's schema concerning dependence (§ 2) and autonomy (§ 3), while applying it to the case of fitness understood as a causal disposition. In doing so, I will consider fitness as a token property (unless otherwise specified). The distinction between type and token is important since it refers to the distinction between *properties* and the *instantiation* of properties, respectively. Properties are not able to enter into causal relations, but instantiations of properties are. Thus, type properties, insofar as they are not instantiated in a particular bearer, cannot enter into

---

[30] The causal autonomy Wilson attributes to weak emergent properties on the basis of their different causal power profiles (see also Wilson 1999; Shoemaker 2001) might be questionable. In fact, I do not agree with this way of conceiving causal autonomy since it seems to imply reduction. I will dwell more on this criticism in § 2.4. In this section, my main interest is merely to introduce Wilson's characterization of both weak and strong ontological emergence.

causal relations. Token properties, conversely, insofar as they refer to the particular instantiation in a bearer of a property of a given type can participate in causal relations (Kim 2003). In this chapter, therefore, I am following Wilson (2016), Chakravartty (2007), as well as Mumford and Anjum (2011) (see Chapter 2, § 3.1) in assuming that the causal efficacy of entities lies in their having efficacious properties.

2. Dependence relations in fitness

As noted above, Wilson distinguishes five kinds of dependence relations: material composition, modal covariation, fusion, non-reductive realization, and causation. In the following sections, I will explain each of them and explore which one is more suitable to capture the relation between individual fitness and its basis.

2.1. Material composition

Material composition is not properly a kind of dependence between properties at different levels, but a characterization regarding the bearers of these properties. According to material composition, there is nothing immaterial (in the sense of an *élan vital*) that characterizes the bearers of properties. Properties can be emergent or not, but the bearer is always a *physical* entity (Stephan 1999, 50).

This criterion of material composition has been considered as a requirement to avoid the physically questionable character that was attributed to the bearers of emergent properties. Since material composition is not a form of dependence but it refers to how the bearers of properties must be, it is compatible with both SE and WE. Furthermore, material composition is also compatible with the absence of emergent properties. This is so since how an entity is (the bearer) is independent of the kind of properties that characterizes it (Wilson 2016).

Concerning fitness, the material composition of organisms, i.e. the bearers of fitness, is widely accepted in contemporary biology. To claim otherwise could imply a commitment to the largely rejected position of vitalism (recall § 1). As previously stated, emergentism is not a form of vitalism since it rejects the presence of immaterial forces in living beings. Exploring the emergent character of fitness, therefore, presupposes the acceptance of the material composition of organisms. Fit organisms are material organisms, and their singularity with regard to non-living entities must rely on their properties instead of their composition.

2.2. Modal covariation

Modal covariation is associated with supervenience. As seen in the previous chapter regarding Rosenberg's supervenient account of fitness (§ 2.1), supervenience is a relation of covariation between two sets of properties, namely the subvenient or lower-level properties, and the supervenient or higher-level ones (Kim 1990). Yet, besides covariation, supervenience has also been characterized as entailing, at least *indirectly*, a relation of dependence (van Gulick 2001). According to Kim (1990), it was Donald Davidson, in his 1970s paper entitled "Mental Events", who characterized supervenience as being more than a mere relation of covariation between two sets of properties. In this regard, besides covariation, supervenience has been considered as also implying a relation of non-reductive dependence, i.e. a relation of dependence that does not entail the reducibility of the supervenient property to its subvenient base:

> Although the position I describe denies there are psychophysical laws, it is consistent with the view that mental characteristics are in some sense *dependent, or supervenient*, on physical characteristics. Such supervenience might be taken to mean that there cannot be two events alike in all physical respects but differing in some mental respects, or that an object cannot alter in some mental respect without altering in some physical respects. Dependence or supervenience of this kind does not entail reducibility through law or definition: if it did, we could reduce moral properties to descriptive, and this there is good reason to believe cannot be done; and we might be able to reduce truth in formal system to syntactical properties, and this we know cannot in general be done (Davidson 1980[1970], 111. Emphasis added).

The recourse to supervenience in order to make sense of emergent properties has been quite common even since the work of the British Emergentists. For instance, C. L. Morgan, one of the theoreticians of the British Emergentism, occasionally used 'supervenience' as a stylistic alternative for 'emergence' (Kim 1990). In the same line, C. D. Broad advocated the idea that "emergent properties must strongly supervene on microstructural properties" (Beckermann 1992, 103; see also Broad 1925; Kim 1990; Horgan 1993). Yet, it was Davidson's characterization of supervenience as a relation of *dependence* without *reduction* that has caused many philosophers to use it in different philosophical fields such as metaphysics and philosophy of mind (Kim 1990; Jackson and Pettit 1990; Horgan 1993). As Terence Horgan has claimed:

> Many philosophers were attracted by the thought that a broadly materialistic metaphysics can eschew reductionism, and supervenience seemed to hold out the promise of being a non-

> reductive inter-level relation that figures centrally in a non-reductive materialism (Horgan 1993, 565).

In metaphysics, authors such as Frank Jackson and Philip Pettit (1990), James van Cleve (1990), O'Connor (1994), Paul McLaughlin (1997) or Chalmers (2006), among others, have also argued for a characterization of emergence in terms of supervenience in order to "ward off the charge that mental (and other) properties being emergent and irreducible are "spooky" properties" (Jackson and Pettit 1990, 145). According to them, to postulate a relation of supervenience between the emergent and the lower-level properties of a system allows making sense of the notion of emergence itself, since supervenience is a form of dependence that does not imply reduction. O'Connor puts this idea in the following way:

> I would argue that supervenience is needed if emergence is to be capable of being incorporated within a scientific framework. Without the two components of determination and dependency, there would be no potential for uncovering precise causal conditions under which emergence occurs (O'Connor 1994, 14).

Characterizing supervenience as a relation of *dependence* makes possible to explain why there is a covariation between subvenient and supervenient properties. The reason is that dependence implies covariation, although the reverse does not necessarily hold. Kim (1990) illustrates this last idea as follows: let's consider three sets of properties *A*, *B* and *C*. The set *A* covaries with the set *B*, whereas the set *B* does not covary with the set *A*. A possible explanation of this situation might be that the set *A* depends on the set *B* whereas the set *B* does not depend on the set *A*. However, Kim argues, this might not be so. It is possible for the set *A* not to depend on the set *B* despite the covariation, insofar as both sets *A* and *B* might, in fact, depend on another set of properties *C*. Covariance, therefore, is not sufficient to imply dependence. For Kim, the concept of dependence is metaphysically richer and deeper than that of covariance. In this regard, it is the characterization of supervenience as implying dependence that allows for the establishment of priority ontological relations between the subvenient and the supervenient properties. Supervenient properties depend on, and therefore, covary with subvenient ones, i.e. subvenient properties are ontologically prior with regard to supervenient properties (Kim 1984, 167).

To claim that there is a dependence relation between two sets of properties, therefore, avoids the possibility of covariation to be caused by another set of properties on which both the supervenient and the subvenient properties might depend (Kim 1984, 1987). However, the understanding of supervenience as a kind of dependence that does not imply reduction is problematic. Kim, for instance, has widely argued that a higher-level property of a system cannot *depend on* the lower-level properties of the parts of the system that instantiates it, *without being reduced* to them (Kim 1984, 1987, 1989, 1990).

According to him, there are two main concepts of supervenience that seem to be at play in philosophical debates, namely i) weak supervenience, according to which "*A weakly supervenes* on *B* if and only if necessarily for any property *F* in *A*, if an object *x* has *F*, then there exists a property *G* in *B* such that *x* has *G*, and if any *y* has *G* it has *F*" (Kim 1984, 163); and ii) strong supervenience, in which "*A strongly supervenes* on *B* just in case, necessarily, for each *x* and each property *F* in *A*, if *x* has *F*, then there is a property *G* in *B* such that *x* has *G*, and *necessarily* if any *y* has *G*, it has *F*" (Kim 1984, 165). Neither of these concepts, notwithstanding, is able to make sense of the idea of dependence without reduction, since either there is a relation of dependence and therefore, it is difficult to justify that there is irreducibility (as per strong supervenience), or there is irreducibility but then, it is not clear that there is a proper relation of dependence taking place between the subvenient and the supervenient properties (as per weak supervenience).

In weak supervenience, the relation of dependence that exists between subvenient and supervenient properties is only guaranteed within a particular given world. Thus, within a particular world, all things that are indiscernible regarding their subvenient properties are also indiscernible regarding their supervenient ones. Since in these cases the correlations between subvenient and supervenient properties lack modal force (Kim 1987, 317), weak supervenience is not able to guarantee stability across worlds. And thus, it does not properly capture a relation of dependence since "determination or dependence is naturally thought of as carrying a certain modal force" (Kim 1984, 160). Due to the lack of modal force, in cases of weak supervenience, reduction of the supervenient property to the subvenient ones does not follow. Yet, without the modal force, it is not possible to explain why there is a connection between certain subvenient properties and certain supervenient ones since, as Kim claims, there seems to be no more than a mere *de facto* coincidence and not a proper relation of dependence between both sets of properties.

Unlike weak supervenience, in strong supervenience there is a modal operator (the second 'necessarily') that guarantees the relation of dependence between the supervenient and the subvenient properties across possible worlds. According to strong supervenience, "for any worlds w_1 and w_2 and any things x and y (in w_1 and w_2 respectively), if x in w_1 is B-indiscernible from y in w_2, then x in w_1 is A-indiscernible from y in w_2" (Kim 1984; Horgan 1993). Yet, Kim argues, if the relation of dependence that characterizes strong supervenience implies modal force, then it is not clear why it is not possible to reduce the supervenient properties to the subvenient ones.

Weak and strong supervenience, therefore, illustrate that supervenience cannot be properly characterized as implying, at the same time, dependence without reduction. As we have seen, in order for supervenient properties to be irreducible to the subvenient ones, the relation of dependence should lack modal necessity, as per weak supervenience. Yet, it is not clear that there is really a relation of dependence in these cases in which there is no modal force (Kim 1984; Beckermann 1992). Conversely, if there is a relation of dependence between the subvenient and the supervenient properties, as per strong supervenience, then the dependence relation will be so strong, insofar as it implies modal force, that it would be difficult to explain why there is no reduction (Kim 1989, 40).

Due to the problems of weak and strong supervenience, some authors have proposed the so-called *global supervenience* as an alternative (Horgan 1982; Lewis 1983a, b). According to Terence Horgan (1982) and David Lewis (1983a, b), global supervenience is able to make sense of the idea of dependence without reduction insofar as the dependence relation is not given between individuals within worlds, but between worlds as wholes. Thus, according to global supervenience: "*A* globally supervenes on *B* just in case worlds that are indiscernible with respect to *B* ("B-indiscernible", for short) are also A-indiscernible" (Kim 1984, 168).

'Indiscernibility' of worlds with regard to a set of properties means that these properties are distributed over their individuals in the same way (Kim 1990). For reasons of simplicity, Kim assumes that the worlds have the same individuals. The key to global supervenience is precisely that there are worlds and not individuals that are compared for discernibility or indiscernibility regarding a set of properties. This implies that, although in a given world there might be covariation regarding the subvenient and supervenient properties, this covariation does not imply a property-to-property correlation between subvenient and supervenient properties, i.e. it does not imply a strong connection between both sets of properties, as per strong supervenience. This is the main reason why it is claimed that global supervenience is able to make sense of the idea of dependence without reduction.

However, global supervenience is problematic as well. Since it does not imply a strong connection between subvenient and supervenient properties, it is not clear that it could imply a relation of dependence at all. In fact, it has been argued that global supervenience does not guarantee a relation of dependence between subvenient and supervenient properties neither in the same nor across possible worlds; that is, global supervenience does not imply either weak or strong supervenience (Kim 1984, 1989, 1990; Petrie 1987; Beckermann 1992). It does not imply weak supervenience since it is possible to conceive of a given world wherein there are two exactly physical duplicates that differ with respect to their supervenient (say mental) properties. Although this case affects weak supervenience, it does no problematize global supervenience since it is still possible to conceive of another world that is indiscernible with this one regarding both the subvenient and the supervenient properties. In the same line, global supervenience does not imply strong supervenience, insofar as in cases in which strong supervenience fails, global supervenience can still be maintained. It is possible to conceive of two different worlds w_1 and w_2, in which there are two different objects, x and y in w_1, and x^* and y^* in w_2, that are alike regarding their subvenient properties but differ regarding the supervenient ones. In this case, although strong supervenience fails, global supervenience is not affected since the two worlds are not completely indiscernible: they differ regarding their supervenient properties. In fact, since it implies worlds indiscernibility, global supervenience allows to conceive of a possible world that although is slightly different from this one with respect to a "minute physical aspect" (such as having one hydrogen atom more) (Kim 1990, 41), it is as different as we pleased regarding the supervenient properties. The existence of this kind of world does not affect global supervenience because the two worlds are not, in fact, indiscernible concerning subvenient properties. Global supervenience, therefore, does not imply reduction, but it does not imply dependence either (Beckermann 1992; Kim 1984, 1989, 1990).

However, and besides the problems of supervenience, Wilson proposes a kind of modal covariation in terms of 'minimally nomological supervenience', according to which "an emergent feature (at least nomologically) requires some base feature, and a given base feature (at least nomologically) necessitates any associated emergent property" (Wilson 2016, 13). According to Wilson, minimal nomological supervenience is compatible with both strong and weak ontological emergence. Regarding strong ontological emergence, when a higher-level property that supervenes on lower-level ones has a new causal power, i.e. a power that is not possessed by any of the lower-level properties of the system, then it is plausible to suppose that there is a nomological connection between both the emergent and the base properties. Concerning weak ontological emergence, it is plausible to claim that the emergent property stands in a metaphysical necessary connection with respect to the base properties of the system in which it is embedded. And, if there is metaphysical necessity, then there is also a nomological necessary connection between the emergent and the base properties.

As previously seen (Chapter 2, § 2.1), some accounts consider fitness as a supervenient property (Rosenberg 1978). In contrast, in CDAF, fitness does not supervenes upon lower-level properties of organisms. Fitness, insofar as it is a disposition, is manifested with a particular intensity, namely the higher or lower ability of an organism to survive and reproduce as compared to other individuals of a population (see Chapter 2, § 4.1). However, this intensity might not vary despite the internal and external perturbations, such as those resulting from diseases and environmental perturbations, that might affect the organism's particular functional dispositions and their respective intensities along its life-history.

The stability of fitness' intensity can be attributed to the fact that fitness is a property of cohesive and robust systems. To claim that a system is cohesive means that there is a connection between the parts that constitute the system such that they are able to move together through space and time making the system itself able to act coherently and resist internal and external fluctuations (Collier and Muller 1998; Collier 1988). Biological organisms are cohesive systems in this sense (Collier and Muller 1998), and this, in some cases, confers them an extreme robustness against internal and external influences. To claim that a system is robust means that it is both persistent and pleonastic. A system is persistent when it is able to guarantee the production of an end-state by compensating changes that occurred within the system itself, and pleonastic when the production of a particular end-state is given by means of alternative pathways (McShea 2012; Austin and Nuño de la Rosa 2018).

Cohesive and robust systems might manifest higher-level properties that are not affected by changes that occur in the properties of the parts that characterize the system. Thus, when some perturbation is affecting some of the properties of the organism, it can be compensated so that no changes in its phenotype are produced (Bertolaso and Cerezo 2017). In this regard, since fitness is a higher-level property of organisms and organisms are cohesive and robust systems, fitness is a property that might not vary in intensity regardless of the changes that might occur in its basis.

I will come back to the notion of cohesiveness in Section 3. The point I want to highlight so far is that, since organisms are cohesive and robust systems, the intensity of fitness might not change despite internal and/or external fluctuations that might affect the properties of the parts upon which it depends. Therefore, we can conclude that since changes in the basis of fitness (i.e. the functional dispositions that characterize the organism) are not associated with changes in fitness, the relation of dependence between them is not one of supervenience.

2.3. Fusion

In order to avoid the problems of defining emergence in terms of supervenience, different authors have approached the notion of emergence in a different way. In this regard, Humphreys has argued that emergent properties depend on the lower-level properties of the parts of the system on which they are instantiated insofar as they result from a 'fusion operation' (represented as [.*.]) among those parts (Humphreys 1997). Humphreys characterizes fusion operation as a process where two lower-level properties (P_m^i and P_n^i) are combined to form a (i+1)-level property [P_m^i * P_n^i].

The fusion operation does not need to be a causal interaction, but can "represent interactions of different kinds" (Humphreys 1997, 10). Yet, in all cases it refers to a "real physical operation", in the sense that it is "neither a logical operation such as conjunction or disjunction nor a mathematical operation such as set formation".

By appealing to the fusion operation, Humphreys rejects the idea that emergent properties are instantiated *at the same time* that a certain lower-level basis is given: after the lower-level properties fuse, they cease to exist and an emergent property appears. The lower-level parts of a system, therefore, go *out of existence* when they give rise to an emergent property (Humphreys 1997). According to Wilson, dependence by means of fusion is compatible with SE. The emergent property results from the fusion operation of the parts that constitute the system, and it is precisely due to this fusion that the emergent property has a *new* causal power (Wilson 2016).

The criterion of fusion operation to recognize emergent properties has been used in different fields such as quantum mechanics (Silberstein and McGeever 1999). Yet, it is not the kind of dependence that seems to characterize the relation between individual fitness and its basis. The different functional dispositions that characterize an organism, i.e. the heart pumping blood, the lungs performing gas exchange between the blood and the air or the liver producing bile, do not fuse and disappear but are simultaneous with fitness itself. In fact, fitness could not be manifested without these functional dispositions operating.[31]

---

[31] Someone might argue that population geneticists fuse the functional properties of organisms when they assign them a single fitness value. However, this fusion is nothing else than a mathematical operation and not a real fusion operation as this criterion implies.

2.4. Non-reductive realization

The core idea of realization is that the function attributed to a higher-level property is, in fact, performed by the lower-level token properties of the system in which it is instantiated (Bennet 2011). Some authors, such as Kim, have appealed to functional realization to argue against ontological emergent properties (see Kim 1993, 1998, 1999, 2006). According to Kim, functional realization implies functional reduction, which consists of three steps (Kim 1999):

1. To attribute a function to the higher-level property we want to reduce.
2. To look for the realizers of that higher-level property in the lower-level, that is, to look for the basic parts or constituents of the system that allow the instantiation of the higher-level property.
3. To elaborate a theory that shows how the realizers of the emergent property have and are able to perform the function that is attributed to the higher-level property.

Functional reduction illustrates that it is possible, at least in principle, to find, in each particular case, the lower-level realizers that perform the function of a higher-level property (but see Needham 2009). Hence, it is not possible to argue for the existence of causally autonomous emergent properties in a system since the higher-level property would inherit *all* the causal powers of its lower-level realizers. Recall Kim's causal inheritance principle presented above (§ 1), according to which "if a functional property *E* is instantiated on a given occasion in virtue of one of its realizers, *Q*, being instantiated, then the causal powers of this instance of *E* are identical with the causal powers of this instance of *Q*" (Kim 1999, 15).

Seen in this light, functional realization is not compatible with emergentism. Since the higher-level property inherits *all* the causal powers of its lower-level realizers, it is nothing over and above them (Wilson 2016). For Kim, therefore, higher-level properties do not have any causal power on their own, since the causal power is realized by the lower-level properties of the parts of the system. The so-called emergent properties, therefore, are nothing more than mere epiphenomena. Wilson, notwithstanding, offers an alternative way of conceiving non-reductive realization that is compatible with emergentism (Wilson 2016). In those cases in which a higher-level property is realized by the lower-level properties of a system without being reducible to them, there is an ontological emergent property. In this case, the higher-level property of a system *does not inherit all* of the causal powers of its realizers, but only a proper subset of them, what Wilson calls the 'Subset of Powers Condition':

> Subset of Powers Condition: Token higher-level feature S has, on a given occasion, a non-empty proper subset of the token powers of the token lower-level feature P on which S synchronically depends, on that occasion (Wilson 2016, 357).

According to Wilson, since the realizing basis has more powers than the higher-level realized property, the latter cannot be reduced to the former (see also Gillet 2002). In this case, therefore, the higher-level property has a *causal profile* that is *different* from that of its basis, which means that it is both causally and ontologically autonomous (see § 1).[32] In other words, the higher-

---

[32] As it will be shortly shown, I do not fully agree with Wilson's characterization of weak emergence in terms of non-reductive realization (2016). In my view, it is not obvious that this

level property is an ontologically weak emergent property. An example of non-reductive realization can be offered in terms of the determinable/determinate relation, such as the relation between colors and shades. Consider that train drivers are said to stop the train when the light of the traffic light is white.[33] White is a determinable that might be multiply determined by different determinates (i.e. the different shades of white) such as 'whitesmoke' or 'snow-white'. Regardless of whether the white of the traffic lights was determined by whitesmoke or snow-white, train drivers should stop the train if they saw the light. In this regard, the different determinates of white share the causal power of the determinable white: advising train drivers they should stop the train. However, they might also have more causal powers besides this one (snow-white might also have the causal power for polar bears to avoid predators, for instance), and they might differ precisely with regard to these other powers. The determinable 'white', therefore, only inherits a proper subset of the causal power of its determinates, which shows that its causal power profile is different from that of its determinates and that it is not reducible to them. In Wilson's words:

> Plausibly, a given determinable will be associated with a distinctive set of powers; moreover, this determinable will typically be 'multiply determined' by associated determinates; distinct determinates of the determinable will share the powers of the determinable, but will differ with respect of their powers (Wilson 2016, 371).

---

form of dependence really implies non-reduction, and that it refers, as a consequence, to a form of emergence: it is precisely due to the fact that the higher-level property inherits its causal power from its basis that it could still be considered as a reducible one.

[33] The following example is used to illustrate the idea of the determinable/determinate relation, and the way in which the determinable is said to inherit a 'proper subset' of the causal powers of the different determinates. I am aware that the determinable white does not have the intrinsic causal power of making trains to stop, but that it is a convention that train drivers have to stop the train when they see the white light in the traffic light.

Is this form of dependence the one that is given in the case of fitness? In other words, is individual fitness non-reductively realized by the particular basis upon which it depends? One way of arguing in this direction is by appealing to its multiple-realizability. Fitness, as a type, is the disposition to survive and reproduce in a particular environment and population. This type of property can be instantiated in different types of realizing bases, i.e. by different combinations of different functional dispositions in different species. This variety of realizing bases might differ due to the different causal powers they might have: birds have functional dispositions that allow them to fly, whereas fishes have others that allow them to swim. Despite their differences concerning their realizing bases, both birds and fishes can be properly said to be fit in their particular environments and populations. Thus, although all types of realizing bases are able to realize the same higher-level property (i.e. fitness), it is not possible to reduce fitness to any of its multiple realizers since the different realizers have more causal powers besides those attributed to fitness. In this sense, fitness, as a token, only has a proper subset of the token powers of the functional dispositions that realize it in each case. If individual fitness is characterized as having merely a proper subset of the causal powers of its basis, then its causal profile would differ from the one of its realizing basis, which means that fitness would be a weak ontological emergent property.

Although this argument is plausible, it does not seem to properly hold in the case of CDAF. The reason is not that fitness is not present in multiple realizing bases (it is), but that it does not meet the Subset of Powers Condition. As previously seen, non-reductive realization implies that in order for a higher-level property to have a proper subset of causal powers, the realizing causal basis must have these causal powers as well. Otherwise, as argued by Wilson, there is nothing the higher-level property can inherit:

> [...] if the distinctive causal role of a realized feature is, on a given occasion, played by a lower-level realizing feature, every token power of the higher-level feature, on that occasion, will be numerically identical with a token power of the feature upon which it synchronically depends, on that occasion (Wilson 2016, 369).

In the case of fitness, it is not clear that the functional dispositions that are the basis of individual fitness have the causal power of making the organism able to survive and reproduce in a particular environment and population. The causal power that fitness should inherit if non-reductively realized does not seem to be found in any of the functional dispositions that characterize the organism. The heart is disposed to (has the causal power of) pump(ing) blood, the legs are disposed to (have the causal power of) make(ing) the organism able to run, and so on and so forth. Since the realizers of fitness do not have the causal power of making the organism able to survive and reproduce in a particular environment and population, it is not clear from where fitness inherits its causal power. The relation of dependence between individual fitness and its basis, therefore, is not one of non-reductive realizability. As I will argue when considering the causal autonomy of fitness (§ 3), the disposition to survive and reproduce is a disposition that belongs to the organism as a whole, rather than its parts. A possible objection to this might be that although the causal power of fitness is not possessed by any of the organism's lower-level functional dispositions, it is possessed by the non-linear combination of the functional dispositions that characterize it. I will come back to this objection in Section 4.

Although for CDAF the form of dependence between fitness and its basis is not the one of non-reducible realization, there are already accounts of fitness in which the causal power of fitness is said to be realized by the properties of organisms. PIF is an example of this (Mills and Beatty 1979). However, in this account, fitness is not considered as a weak ontological emergent property (see Chapter 2, § 2.2). Instead, it ends up being reduced to the particular traits and functions (i.e. the functional realizers) that are considered to be responsible for the actual survival and reproduction of the organism in each case (Sober 1993). In other words, PIF reduces individual fitness to the particular anatomical, physiological or behavioral traits of an organism, all of which contribute to its viability to leave offspring in a particular environment.

I think that the reductionist conclusion of PIF regarding fitness follows precisely from appealing to functional realizers in order to explain the causal power of a higher-level property. Despite the fact that the realizing basis is said to have more powers than those inherited by the realized property, the inherited powers *are still perceived as being performed by the realizers*, and thus, I consider that there is no reason why the causal autonomy of the realized property needs to be maintained.

2.5. Causation

Another way the relation of dependence between an emergent property and its basis is understood is by means of causation (O'Connor and Wong 2005; Wilson 2002, 2016). For O'Connor and Hong Yu Wong, the relation that is given between the emergent property of a system and the lower-level properties of the parts of the system it depends on is causal: the lower-level properties of the parts of the system *cause* the emergent property to appear:

> Emergent properties are basic properties, *token*-distinct in character and propensity from any microphysically structured properties of their bearers. If their appearance in certain systems is to be explained at all, they must be explained in terms of *causal*, not purely formal, relationship to underlying, immediately *preceding* structures (O'Connor and Wong 2005, 664).

Causation is a kind of dependence compatible with SE, insofar as effects are said to have powers that are different from their causes (Wilson 2016): the effect is not identical to the cause. Concerning fitness as a causal disposition, I claim that the kind of connection that links individual fitness with its basis is a causal one. This idea is already implied in CDAF since it is understood as the resultant (i.e. effect) of the non-linear combination of functional dispositions that characterize the organism. However, in Chapter 2, I have mainly paid attention to the characterization of fitness as well as to how this way of understanding it allows to solve and clarify some of the problems that have arisen in the philosophy of biology concerning the ontological characterization of this property. In this section, notwithstanding, I will focus on explaining in more detail the causal process (and its features) that gives rise to individual fitness.

As we have previously seen (Chapter 2, § 3.2), according to DTC, causation is a process that takes place once a disposition has been triggered and the particular manifestation associated with it is being produced. Understood in this sense, the causal process is mainly characterized by i) *simultaneity* -both causes and effect entirely coincide temporally; ii) *non-necessitarianism* -causes tend towards their effects without necessitating them;

and iii) *compositional pluralism* -causes do not act in isolation but they work (either linearly or non-linearly) in combination.

These three main features of DTC also hold in the causal process that gives rise to fitness. In this case, the functional dispositions that characterize an organism are the basis of its fitness. These functional dispositions, insofar as they are manifesting (e.g. the heart is pumping blood), are acting as *causes.* Furthermore, these causes do not work in isolation but they are *non-linearly combined,* insofar as they need each other in order to properly develop and perform their tasks (Gilbert and Sarkar 2000). In this regard, functional dispositions are continuously affecting each other such that the way a particular trait performs its task can affect how a different trait performs its own. As a consequence, it is not possible to treat or consider a particular function and its effect without taking into account the other functions that characterize the organism and their own effects (Collier and Muller 1998).

This non-linear interaction among functional dispositions is illustrated by some pathologies. Some pulmonary diseases, for instance, can affect the way lungs perform their function. Yet, they can also alter the way the heart performs its function and the effect it produces due to the connection that exists between the lungs and the heart as performed by the pulmonary arteries. The chronic obstructive pulmonary disease is an example of this, although all chronic pulmonary diseases can affect the function of the heart. In the same way, a malfunction, or a failure in the function of the heart (it does not pump blood adequately due to an arrhythmia, for instance), can alter the way lungs perform their activities due to an accumulation of blood in the lungs, for example.

Insofar as functional dispositions are *non-linearly* combined, they are able to produce an effect that is more than a mere addition of the different effects they produce in isolation. This effect, as I am considering it, is individual fitness. In other words, it is in their non-linear combination that functional dispositions engage in a process in which a *fit organism* appears as the *simultaneous* effect. Notice that, in this sense, all living organisms, insofar as resulting from the combination of functional dispositions, are fit organisms. Another question refers to the particular level of fitness each organism has. Regarding this point, it is also important to recall the distinction between fitness and expected fitness (Chapter 2, § 4.3). Fitness is the organismic property that results from the non-linear combination of functional dispositions that characterize the organism. This is the ontological way of characterizing fitness. Conceived in this way, all living organisms are fit organisms. However, not all organisms have the same level of fitness. The level of fitness, as previously explained, is established by paying attention to the expected fitness each organism is calculated to leave, and it might be different from the number of offspring it actually leaves.

Finally, another feature of the causal process that gives rise to fitness, and that it is in line with the framework of DTC, is *non-necessitarianism*. There is not a necessary connection between individual fitness and its basis. As argued in Chapter 2 (§ 4.3.2), preventing factors might affect the causal process that gives rise to fitness such that it stops manifesting with the result of killing the organism. Furthermore, interfering factors might also intervene in the causal process such that the fit organism might be different from the one that would have resulted without the interfering factors. In this case, it is possible for the organism to have a higher or lower fitness intensity than the one it would have had without the interfering factors.

As stated above, in CDAF, insofar as conceived of from the metaphysical framework of DTC, the relation between fitness and its basis is causal. Here, I have just introduced how the features attributed to causation by DTC are given in the particular case of fitness. Due to the simultaneity between fitness and its causal basis, the synchronic condition of emergentism is fulfilled in this case (recall the definitions of SE and WE in § 1). Concerning the ontological status of fitness as a strong or weak emergent property, causation is compatible with SE. Effects are different from their causes insofar as they have *powers* different from them (Wilson 2016, 368). Fitness, therefore, might be compatible with SE. However, something more needs to be said regarding fitness in order to properly claim that it is a strong emergent property. In particular, it needs to be argued that the causal power of individual fitness is different from that of its basis, i.e. that fitness is a causally autonomous property. As previously stated (§ 1), causal autonomy implies ontological autonomy, although the reverse might not hold. If the causal power of fitness is not autonomous, then either fitness is equated to the functional dispositions that characterize the organism (as per in PIF), or it is a mere epiphenomenon, i.e. a real ontological property that does not have any causal power on its own (as Rosenberg might argue). In the following section, I will argue for the causal autonomy of fitness.

3. The autonomy of fitness

Initially, in this chapter (§ 1), I have characterized the notion of causal power in the same line as Wilson's (2002). In this view, having a causal power means that the fact that a property is instantiated entails that its bearer will have the power, in appropriate circumstances, to cause some effects or intervene in the causation of some effects. In order to claim that fitness has causal power, therefore, it is necessary to show that it allows its bearer to cause some effects.

In § 2.4, I have rejected the idea that the functional dispositions that characterize an organism are the bearers of fitness: neither of them has the causal power of making the organism able to survive and reproduce in a particular environment and population. Furthermore, I have also claimed (§ 2.5) that it is the fit organism that results from the non-linear combination of its functional dispositions that has such power. In doing so, I am assuming the organism to be the bearer of fitness. Yet, before arguing for the causal autonomy of fitness (§ 3.2), something more needs to be said concerning the ontological status of organisms and their features in order to make sense of the idea that they are the bearers of fitness (§ 3.1).

3.1 Conceptualization of organisms from the philosophy of biology

Insofar as in CDAF organisms result from a non-linear causal process, they need to be conceived of as something more than the mere sum or combination of their functional dispositions. The ontological autonomy of the organism, *as a whole,* is the major tenet of organicism (see Collier 1988; Collier and Muller 1998; Kauffman and Clayton 2006; Boogerd et al. 2007; Witherington 2011), which has experienced a recent revival in biological fields such as systems biology, evo-devo, or developmental biology, and that defends both the ontological status of organisms as well as their explanatory role:

> [In organicism] complex wholes are inherently greater than the sum of their parts in the sense that the properties of each part are dependent upon the context of the part within the whole in which they operate. Thus, when we try to explain how the whole system behaves, we have to talk about the context of the whole and cannot get away talking only about the parts (Gilbert and Sarkar 2000, 1).

> The whole cannot be fully understood through decomposition into temporally based lines of part-to-part relations, for such relations gain meaning through their very embeddedness within the whole. Wholes are not simply something to be explained but in fact serve as an explanatory context for their parts. The very nature of the task itself and the functional significance of an organism's action in the task context derive meaning from that organism's particular level of developmental organization (Witherington 2011, 85).

In philosophy of biology, organisms have been conceptualized as *organizationally closed and differentiated self-maintaining systems* (Mossio et al. 2009; Witherington 2011; for an overview of the work developed in this line by the San Sebastian IAS-Research group, see Moreno and Mossio 2015). Self-maintaining systems are systems whose existence depends on the mutual causal relations that are given between the macroscopic pattern of the system itself and the microscopic dynamics of parts that give rise to the system. This idea is usually illustrated by means of a candle flame, where the microscopic reactions of combustion generate the macroscopic pattern (the flame itself), while the flame allows the maintenance of the combustion reactions that guarantee its own existence.

Biological systems are self-maintaining systems in that sense. Furthermore, they are *organizationally closed* and *differentiated* self-maintaining systems. They are *organizationally closed*, insofar as there is a circular causal relation between the activity of the parts of the system (that contribute to the existence of the system itself) and the activities of the system (which are also necessary for its own maintenance, insofar as they contribute to the existence of the activities of its parts) (Mossiot et al. 2009, 824). Biological systems are also *organizationally differentiated,* insofar as the system is able to create distinctive structures "that make a specific contribution to the conditions of existence of the whole organization" (Mossio et al. 2009, 826). In characterizing biological systems as "organizationally closed and differentiated self-maintaining systems", the primacy of the whole living system over its parts is acknowledged. In fact, it is considered that the parts themselves cannot be properly understood without taking into account the system they compose (Witherington 2011, 77)

In the same line as organicism, in CDAF organisms are ontologically conceived as being more than the combination of their functional dispositions. This is an important aspect to consider regarding the emergent character of fitness. On the one hand, because emergent properties have a place within organicism (Gilbert and Sarkar 2000, 2). And, on the other hand, because emergent properties are systemic properties, i.e. properties that cannot be possessed by the parts that constitute a system, but by the system itself (Humphreys 1997; Stephan 2002; Van Gulick 2001).

It is important to keep in mind that being a systemic property is not the same as being emergent. A *systemic property* is a property of the whole, while a *systemic emergent property* is a property of the whole that is also causally autonomous (recall § 1). Thus, although all emergent properties are systemic properties, not all systemic properties are emergent properties. In order to be emergent, a systemic property must have a new causal power (as per strong ontological emergent) or a different causal power profile (as per weak ontological emergence). By highlighting the systemic character of the living organization, therefore, I am not arguing for the emergent character of fitness. Instead, what I want to illustrate by introducing organicism is the possibility of conceiving organisms as the bearers of fitness, that is, as the ones that are able to enter into causing some effects insofar as they are fit.

Besides being considered as "organizationally closed and differentiated self-maintaining systems", in philosophy of biology organisms are also conceptualized as having *natural agency*, insofar as they are able to *interact* with their environment and population (Moreno and Barandiaran 2004; Barandiaran and Moreno 2008). Interaction differs from reaction. The functional dispositions that characterize an organism might *react* to different external or internal stimuli by producing a particular response. The skin, for instance, reacts to sun exposure: the ultraviolet rays stimulate melanocytes, which produce melanosomes that favor the thickening and make the skin to get darker. Organisms, notwithstanding, are not considered to merely react to the different stimuli they might face, but they interact with their environment and population insofar as they perform certain activities or behave in certain ways that are oriented to ensure their survival (Moreno and Barandiaran 2004, 18), and reproduction. In this sense, the skin reaction to sun exposure does not take place for the sake of the survival nor the reproduction of the organism. Rather, it is the organism, as a whole, who avoids or looks for sun exposure depending on whether it needs it for its survival or reproduction.

As I see it, the natural agency that is attributed to organisms is grounded on fitness. Organisms, as previously stated, are the bearers of fitness. And fitness is a property that allows its bearer to behave in such a way that it guarantees its survival and reproduction. In this regard, organisms are natural agents, i.e., they behave in a way that ensures their survival and reproduction, insofar as they are fit. To claim that natural agency is grounded on fitness might seem to imply some form of teleology or appealing to final causes: the organism interacts with its environment and population as it does *in order to* guarantee its survival and reproduction. This teleology, notwithstanding, does not cause real problem insofar as it is natural teleology. In fact, it is implied in the same notion of agency: "[agency] requires the framing of living phenomena in final causal, teleological terms, wherein 'purposiveness' or 'that for the sake of which' provides a key component of scientific explanation" (Witherington 2011, 80).

As I am conceiving it in this section, fitness is a property of whole organisms that makes them able to survive and reproduce in a particular environment and population. And it is on the basis of their fitness that organisms can act as natural agents that interact with their environment and population in order to guarantee their survival and reproduction. Yet, as previously stated, a property might have a causal power and be a systemic one without being emergent, by being a mere epiphenomenon, for instance. In order to be emergent, a systemic property might be causally autonomous. I will argue for the causal autonomy of fitness in the following section.

3.2 The causal autonomy of fitness

The notion of causal autonomy might be conceived of in two different forms depending on whether the emergent property is strong or weak. If a higher-level property is strongly emergent, then it has *at least one token power not identical with any token power of P on that occasion*. Conversely, if it is weakly emergent, then it *has a non-empty proper subset of the token powers had by P, on that occasion* (Wilson 2016, § 1). Concerning fitness, its compatibility with SE has been highlighted by means of its causal connection to its basis. In this sense, I will claim that fitness is a strong emergent property of organisms by arguing that it introduces a new causal power into the world, one that is downwardly exerted.

Recall the consideration of downward causation as the hallmark of emergentism (§ 1):

> There is no question that emergentists should want downward causation. Emergent properties must do some serious causal work, and this includes their capacity for projecting causal influence downward (Kim 2006, 198).

If fitness can *downwardly affect* the functional dispositions that characterize an organism in order to guarantee its survival and reproduction, then it has a new and autonomous causal power, i.e. a power that is not realized by the combination of its functional dispositions. In this regard, I will argue that an organism, insofar as it is fit, can *downwardly affect* the biological functions that compose it in order to favor its own survival and reproduction in a given environment and population.

Downward causation concerning survival

Regarding the survival dimension of fitness, I argue that the way an organism responds to cases of malfunction should be interpreted as an instantiation of fitness' downward causation. In the organizational approach, organisms are the ones responsible for maintaining their own existence by means of creating, constraining, and maintaining the parts that compose them (Mossio et al. 2009; Witherington 2011; Moreno and Mossio 2015). In some cases, the components of the organism might be affected such that they do not perform the function they are supposed to do and that might contribute to the existence of the organism itself. In these cases, the organism can produce new constraints over the other parts and functions of the system such that the malfunction can be compensated and it can guarantee its own survival.

The changes involved in phenotypic plasticity, namely the capacity of organisms to develop several phenotypic outcomes depending on environmental factors, are a good illustration of the downward causal power of fitness. This idea can be illustrated with one of the examples of plasticity discussed by Mary Jane West-Eberhard, namely the case of the goat that was born with a congenital defect that impeded her to walk on its four legs (West-Eberhard 2005). The goat learned to walk on its hind legs, giving rise to a reorganisation of the muscles and bones of the limbs, "including a thickened and elongated gluteal tongue and an innovative arrangement of small tendons, a modified shape of the thoracic skeleton, and extensive modifications of the pelvis" (West-Eberhard 2005, 6545). In other words, the goat was able to modify the form and function of its hind limbs such that they played the same role as if she had four limbs.

I consider that it is due to the fitness of organisms that they are able to downwardly affect the parts that compose them. The organism's tendency to survive and reproduce, and to do so in better conditions (i.e. to be fitter), makes it able, as a whole, to constrain the form and function of its lower-level parts such that it can continue existing. And this constraint exerted by the organism seems to be better explained in terms of downward causation. It is due to the organism's fitness that it can downwardly affect its lower-level parts in order to guarantee its survival and reproduction. As Witherington claims:

> Through downward causation, systems qua systems coordinate and control the properties of their lower-order components and the interactions of those components with one another, preserving their very organization despite flux, renewal and turnover in the lower-order components that comprise them (Witherington 2011, 75).

As previously stated (§ 1), the causal efficacy of entities lies in their having efficacious properties. In this sense, if the organism were not 'fit' (i.e.

if it did not possess the property of fitness), then it would not have a capacity to survive, and no change or constraint on its lower level parts would be made in order to guarantee its survival. In this kind of cases, it seems plausible to claim that fitness allows the organism to achieve its survival threshold by allowing it to affect the parts that compose it. The causal power of fitness, therefore, is manifested in a downward manner.

Downward causation concerning reproduction

In order to illustrate the downward causal power of fitness concerning reproduction, I will appeal to two biological strategies performed in different species, namely delayed reproduction and tactical deception.

Some female organisms have a reproductive strategy called *reproductive delay* that allows them to arrest reproduction until suitable environmental conditions are met (Orr and Zuk 2014). Reproductive delay is common in insects but also in mammals such as bats *(Chiroptera)*, or armadillos *(Dasypodidae)*. It can be manifested in three different ways: *i) delayed fertilization* -females arrest fertilization by delaying the ovulation period or storing the sperm of different males; *ii) delayed implantation* -the blastocyst is not implanted in the uterus but remains unattached in the female reproductive tract (McGowen et al. 2014); and *iii) delayed development* -the development of the embryo is arrested.

These kinds of reproductive delay are mainly manifested in cases of environmental uncertainty (Koons et al. 2008), that is, in cases wherein the environmental conditions are not suitable for the organism to leave offspring due to climate conditions, predators, or lack of food resources. Sometimes, reproduction could be arrested for an entire season, or even a year, e.g.

Rhinolophus landeri. In cases of environmental uncertainty, the female delays reproduction until there are more resources, better climate conditions or fewer predators. Furthermore, the female can also delay reproduction due to populational factors. In this regard, some females can arrest fertilization by storing the sperm of different males until there are not available males in a population, or in cases wherein there has been no opportunity for pre-copulatory mate selection and the female stores the different sperms in order to lately choose among them. This is called 'cryptic female choice' (Orr and Zuk 2014), and occurs in different species such as Drosophila flies and field crickets.

Reproductive delay, therefore, is a strategy that allows the organism to have more offspring than in those cases where reproduction is not arrested (Koons et al. 2008). Different interpretations of this phenomenon are offered, such as its understanding as a mere physiological response of organisms to environmental conditions. In this regard, some authors have characterized it as a response to lower body temperatures (Hamlett 1935). Yet, this interpretation has been discarded on the basis that many taxa that have reproductive delays live in tropical or warm habitats, e.g., *Artibeus jamaicensis*, or *Macroglossus minimus* (Orr and Zuk 2014). As I conceive it, it is due to fitness that some organisms are able to arrest reproduction until the right circumstances are given. If organisms were not fit in this sense, then they would reproduce in a reactive way. That is, they would reproduce in any circumstance and with negative consequences for their offspring, such as being damaged or even killed. However, this is not the case. Organisms do not merely react to the conditions they are embedded in, but they interact with their environment and population, such that they are 'aware of' the accurate or not accurate conditions for having offspring or not. Thus, I consider that it is due to its fitness that the organism is able to reproduce, and delay reproduction is a form of guaranteeing reproduction by downwardly affecting the mechanism that can arrest reproduction. Fitness, in this sense, is manifesting its causal power in a downward manner.

As previously claimed, *tactical deception* is a further example that can help to illustrate the way fitness downwardly exerts its causal power. In some species of cuttlefishes, such as *Sepia plangon*, males employ different strategies to guarantee their reproduction, such as display mate guarding, displace rivals, or interrupt courtship attempts. Recently, it has been argued that males also use a tactical deception strategy for mating a female (Brown et al. 2012). In particular, it has been observed that some males mimic a female display towards their rival males on one side of their body "while simultaneously displaying typical male courtship patterns towards potential mates on the other side. Thus, the courting male is able to perform courtship without being interrupted" (Brown et al. 2012, 729).

This deceptive dual strategy allows the organism to have a higher number of offspring in those cases in which reproduction does actually occur by means of using this deceptive strategy. This kind of tactical deception is usually performed in those environments wherein there are only one more male competitor and a female. In other cases, such as those wherein there are two male competitors and one female, it is more difficult for the courting male to properly orient itself between the female and the other two male competitors.

In explaining tactical deception, one might resort to some kind of signal in the environment (e.g the fact that there are only one male competitor and one female) that activates this mechanism for a deceptive signal such that courtship in cuttlefishes is possible. However, as in the case of delay reproduction, I don’t agree with this interpretation. Tactical deception might be really costly for the courting male since signals among organisms of the same species are said to be inherently honest (Guilford and Dawkins 1991). Thus, not all courting male cuttlefishes use this strategy to mate a female despite the circumstances being the accurate ones. As Culum Brown, Martin Garwood and Jane Williamson claim:

> The proportion of cheats in a population generally remains low, because the cheat’s bluff is occasionally called and the costs of being caught cheating may be considerable (e.g. if the cheater is forced to fight) (Brown et al. 2012, 729).

If it were the environmental factors that activated the lower-level mechanism for deceptive dual display, then the mechanism would be activated in all situations in which the environment was the accurate one. Yet, this is not the case. Just like in delayed reproduction, the disposition to reproduce is not merely triggered by external factors, but internally and contextually regulated. Another objection to this idea might be that this kind of deceptive mechanism is only activated in some courting males but not in others, insofar as not all of them possess this mechanism. I believe this is not the way this phenomenon should be interpreted either. The reason is that some males that have already performed deceptive dual display to mate in previous circumstances, might not perform it in other situations in which the environmental conditions also hold. It has been shown that the more the courting male uses this strategy, the more possibilities it has to be discovered by other males, and this might be highly costly for him (Guilford and Dawkins 1991; Brown et al. 2012).

In some sense, therefore, it can be claimed that it is the fit organism that interacts with its environment and population and activates or not this deceptive mechanism when the appropriate circumstances are given in order to reproduce. Brown et al. put this idea as follows:

> [Courting males] *recognize* when only one rival male is present and only then *elect* to employ a deceptive display. Courting males may *refrain* from such displays if they *detect* more than one rival male perhaps because there is a high probability that their deception will be discovered, the tactic will fail and they will be punished by larger males (Brown et al. 2012, 730. Emphasis added).

I consider that it is due to its fitness that the organism is able to interact with its environment and population in order to guarantee its survival and reproduction. If the organism did not have this tendency, then it would not activate this deceptive signal in order to increase its opportunities to mate and reproduce. Fitness, therefore, manifests its causal power in a downward way in this kind of cases as well.

***

In this section, it has been argued that the capacity of organisms to survive and reproduce in a particular environment and population refers to a causal power they possess due to their fitness. This causal power is an autonomous one, insofar as it is a new causal power that it is downwardly exerted. Due to its causal connection to its basis (§ 2.5) as well as its autonomous character (§ 3.2), fitness can be characterized as a strong emergent property of organisms.

4. Considerations on overdetermination

In the two previous sections, I have argued for the emergent character of fitness on the basis of the causal connection between organismal fitness and its basis, as well as the causal autonomy of fitness. Causal autonomy is a key aspect for a property to be emergent since, otherwise, it would not introduce something new into the world, and thus, it would be a mere epiphenomenon. In this respect, I have claimed that the causal power of fitness is not performed by any of the functional dispositions of an organism alone (§ 2.4). Yet, someone might claim that the causal power of fitness, i.e. making the organism able to survive and reproduce in a particular environment and population is, in fact, performed by the *non-linear combination* of those functional dispositions. If this is so, then, following Kim's causal exclusion argument (§ 1), fitness is not a causal autonomous property, which implies that it is not an emergent property either. If the causal power attributed to fitness is already possessed by the non-linear combination of functional dispositions that gives rise to it, then we should reject the causal autonomy of fitness in order to avoid overdetermination.

In this section, I argue that there is no overdetermination in the case of fitness, insofar as both the non-linear combination of functional dispositions and fitness are, in fact, different 'components' of two different, although 'temporally overlapping', causal processes. Recall here that temporally overlapping causal processes are those in which a particular process (CP_2) takes place insofar as the resultant effect of a different current causal process (CP_1) has been triggered. In these cases, we have two overlapping causal processes since both do temporally coincide and it is the effect of one of them which allows initiating the other once the triggering conditions are met (see Chapter 2, § 3,1).

Two temporally overlapping causal processes are required in the case of fitness: CP_1, which gives rise to fitness, and CP_2 that starts once fitness is triggered and that gives rise to the survival and reproduction of the organism. In this sense, overdetermination is not implied since both causal processes are producing different effects. Concerning CP_1, the causal process consists of the non-linear combination of the functional dispositions that characterize an organism (§ 2.5. See also Chapter 2). Functional dispositions, insofar as they are triggered, are acting as causes and producing their own manifestations (e.g. the heart is pumping blood). These functional dispositions, as previously explained (§ 3), are non-linearly combined such that they give rise to a causal process in which fitness, i.e. the fit organism, is the resultant effect.

Yet, the resultant fit organism is not given in isolation, but it is already embedded in an environment and population. These factors allow the second causal process involved in fitness (CP_2) to start. There are the environmental and populational factors that trigger an organism's fitness such that its capacity to survive and reproduce is manifesting. Survival and reproduction are the effects of the organism's having fitness, and they are manifesting while functional dispositions are non-linearly combined and causally producing the fit organism itself.

One might counter-argue that CP_2 is not necessary since CP_1 already causes the effect of an organism's survival and reproduction. However, I do not consider this to be the case for the reasons alluded in the previous section. As stated above (§ 3.1), functional dispositions are not able to *interact* with the environment and population the organism is embedded in. Instead, they merely react to those external factors. It is the organism, insofar as it is fit, that interacts with its environment and population in a certain way in order to guarantee its own survival and reproduction. In this regard, it seems to me that if fitness were not something different from the non-linear combination of the organism's functional dispositions, then cases in which the organism survives and reproduces by downwardly affecting some of the functional dispositions that characterize it would not be accurately explained. Again, it is the organism tendency to survive and reproduce that makes it able to affect the functional dispositions that compose it in a certain way when the environmental and populational conditions require it.

5. Conclusion

In this chapter, I have explored the ontological status of fitness, and in particular, whether it is a case of an emergent property. In doing so, I have paid attention to two conditions of emergent properties, namely dependence and autonomy, in order to see whether and how they are given in the case of fitness. Regarding dependence, I have argued that the kind of connection that exists between fitness and its basis is a causal one (§ 2). Concerning autonomy, I have argued that the causal autonomy of fitness can be justified by showing that it allows its bearer, i.e. the whole organism, to exert downward causal influence on the functional dispositions that characterize it in order to guarantee its survival and reproduction (§ 3). Since the causal connection between fitness and its basis is associated with strong ontological emergence, and since the causal autonomy of fitness has also been argued against possible objections (§ 4), I conclude that fitness can be characterized as a strong ontological emergent property of organisms. This ontological characterization of fitness might have epistemological implications such as irreducibility or unpredictability. However, my aim here is merely ontological, and thus, I leave open the study of the epistemological consequences that can follow from the ontological nature of fitness.

Chapter 4:
Parthood Relations in Holobionts

> *"All animals and plants establish symbiotic relationships with microorganisms. Often the number of microorganisms and their combined genetic information far exceeds that of their host cells. We refer to the host and its symbiont population as the holobiont, and the host genome and the genomes of all the symbiotic microorganisms as the hologenome."*
> (Rosenberg et al. 2007)

In philosophy of biology, there is a wide debate concerning the characterization of biological individuals (Wilson 2007). The so-called 'Problem of Biological individuality' refers to different questions such as what makes something a biological individual? what is the nature of being a biological individual? or what is the best explicative definition of biological individual? (Clarke 2010, 2012). Generally, biological individuals have been equated to organisms, which is called *the organism-centered view* (Wilson and Barker 2013). However, this view has been challenged on the basis that there seem to be different types of biological individuals (e.g., evolutionary, developmental, metabolic, and others). Furthermore, during the last decade of the 20th century, discoveries concerning symbiotic relations given in different species such as Coral reefs and fungus, as well as the realization of the near omnipresence of this phenomenon, have also affected the way biological individuality is conceived. In this context, the notion of 'holobiont' has been recently coined to refer to the symbiotic assemblages composed by a macroorganismic host plus its symbiotic microbiota. Following the usual terminology in the field, I will refer to the host as the 'macrobe' or 'macro-organism' because, by default, the host will always be a multicellular organism (O'Malley and Dupré 2007). In turn, the microbiota refers to "[t]he assemblage of microorganisms present in a defined environment" (Marchesi and Ravel 2015, 31).

In recent years, the ontological status of holobionts has become a hot topic in philosophy of biology. It has created a great amount of controversy with regard to whether holobionts are real biological individuals (the so-called 'holobiont hypothesis') or just ecological communities composed of independent entities that interact together (a position that can be labeled 'the ecological-community view'). Lynn Chiu and Gérard Eberl (2016) have recently developed a middle-ground position between these two alternative views. According to what they call *the eco-immunity account of the holobiont*, holobionts are not biological individuals nor ecological communities but hybrids between a host and its microbiota. The authors base their claim on different arguments that reject the microbiota to be a proper part of the host. Insofar as this account directly appeals to the notion of parthood, I consider that metaphysical considerations regarding mereology can be relevant for, as well as shed some light on the philosophical debates concerning biological individuality.

In this chapter, I pay attention to Chiu and Eberl's account and explore some of its metaphysical commitments and implications. In particular, I focus on the characterization of 'parthood' that grounds this account in order to make explicit some of the metaphysical implications that follow from it. In doing so, I will argue that Chiu and Eberl's characterization of parthood implies counterintuitive conclusions regarding what should be considered as a proper part of the host.[34] As a consequence, this account does not support the claim that the microorganisms of a host's microbiota are not proper parts of the host, at least with regard to resident microbiota.

The structure of the chapter is as follows. Firstly, I briefly introduce the debate concerning the ontological status of holobionts (§ 1). Then, I develop in more detail the eco-immunity account (§ 2). In the last section (§ 3), I focus on how the notion of parthood is conceived within the eco-immunity account and present some of the counterintuitive conclusions regarding parthood relations that follow from it. I divide this section into three parts in which I argue that: 1) causal-functional integration is a kind of criterion for parthood in microbiota (§ 3.1); 2) proper parts of an object can change (§ 3.2), and 3) parthood is not a special case of scaffolding (§ 3.3).

1. The debate on the ontological status of holobionts

The current realization that multicellular forms of life require of an abundant number of microorganisms to survive and perform their basic biological functions has led some biologists and philosophers of biology to argue for the holobiont as a new level of biological individuality (Margulis and Fester 1991; Zilber-Rosenberg and Rosenberg 2008; Dupré and O'Malley 2009; McFall-Ngai et al. 2013; Rosenberg and Zilber-Rosenberg 2014, 2016; Bordenstein and Theis 2015; Theis et al. 2016). Holobionts, in their simplest form, are the biological entities that result from the symbiotic association between a given host and its symbiotic microbes, "including those which affect the holobiont's phenotype and have coevolved with the host (…), those which affect the holobiont's phenotype but have not coevolved with the host (…), and those

---

[34] By talking of 'proper parts' the possibility of a whole to be a part of itself is discarded. A proper part of an object, therefore, is one which is not identical with the whole itself.

which do not affect the holobiont's phenotype at all" (Theis et al. 2016, Fig 1). The role of the host[35] is played by what Maureen O'Malley and John Dupré have characterized as a 'macrobe' (animal, plant), i.e. a visible entity with a considerable degree of cellular and morphological sophistication (O'Malley and Dupré 2007). In turn, the microbes can be very diverse, including *Archaea*, eubacteria, fungi, or viruses (Rosenberg and Zilber-Rosenberg 2014, 2016).

Regarding the relation between microorganisms and macroorganisms, some observations have been made in current microbiology that illustrate firstly, that symbiosis is a universal phenomenon in the living world: all animals and plants engage in constant symbiotic relationships with microbes of different species; secondly, macrobes are dependent on their symbiotic microbes to realize their more basic biological functions, including their development and metabolism; and thirdly, the intervention on the normal microbiota that interacts with a given macrobe leads to an altered system, sometimes generating diseases (Reshef et al. 2006; Rosenberg et al. 2007; Gordon et al. 2013).

Although observations concerning the interaction between microorganisms and macroorganisms were originally made about corals, the advocates of the holobiont concept claim that, in fact, it is not possible to understand biological individuals independently of the microbiota they interact with. This has led to the formulation of the so-called *holobiont hypothesis,* according to which all animals and plants have to be considered as holobionts (host *plus* its symbiotic microbiota), which are a genuine new level of biological individuality, as well as a unit of selection in evolution

---

[35] By 'host' I will always mean the monogenomic individual derived from a zygote or from an asexual division.

(Rosenberg et al. 2007; Zilber-Rosenberg and Rosenberg 2008; Gordon et al. 2013).

According to the advocates of the holobiont hypothesis, those observations concerning holobionts jeopardize some of our most traditional intuitions about the biological individuality of the macrobes (Gilbert et al. 2012; Stencel and Proszewska 2018).[36] The biological individual is not anymore the genetically homogeneous macrobe that develops from a zygote, but a multispecies and polygenomic entity –the holobiont itself (Dupré 2010; Rosenberg and Zilber-Rosenberg 2014). In this regard, they argue, holobionts are biological individuals in an anatomic, physiological, developmental, and even in an evolutionary sense (Rosenberg et al. 2007; Zilber-Rosenberg and Rosenberg 2008; Gilbert et al. 2012; Gordon et al. 2013; Rosenberg and Zilber-Rosenberg 2014, 2016; Bordenstein and Theis 2015; Shropshire and Bordenstein 2016; Theis et al. 2016; Lloyd 2017; Roughgarden et al. 2017).

Nonetheless, this strong thesis about holobionts has been recently challenged by different authors on the basis of two major arguments. Firstly, the hypothesis that the holobiont is a biological individual is too vague and it is not precise enough to be biologically significant (Booth 2014; Skillings 2016; Queller and Strassmann 2016; Chiu and Eberl 2016). Secondly, the claims, assumptions, and implications concerning the biological individuality of holobionts do not follow from current empirical evidence, or it is not properly based on it. In general, the advocates of *the ecological-community view* claim that the current empirical evidence is not sufficient to claim that holobionts are functionally integrated entities, and therefore, any claim about their status as biological individuals is ungrounded. Some of the assumptions made by

---

[36] 'Individuals' in this context, should be understood as meaning "genetic homogeneous individuals" (Dupré 2010).

the holobiont hypothesis advocates about the types of interactions that tie host and microbiota together only apply to some members of the microbiota, but they do not apply to the microbiota as a whole, which questions the general consideration of the holobiont as a unified biological individual (Skillings 2016; Moran and Sloan 2015; Douglas and Werren 2016; O'Malley 2017; Hurst 2017).

In this regard, it is important to note that detractors of holobionts' individuality do not deny that some members of the microbiota might be sufficiently integrated with the host so as to constitute a unique biological individual (e.g. mitochondria in eukaryotes, *Buchnera aphidicola* in aphids, *Wolbachia pipientis* in some insects, etc.). What they deny is that a host plus *all* its symbiotic microbiota act together as a unique biological individual, and they find more accurate to talk about independent individuals that live together in association (Booth 2014; Moran and Sloan 2015; Godfrey-Smith 2015; Chiu and Eberl 2016; Douglas and Werren 2016; Hurst 2017). The same position has been defended from an evolutionary perspective, claiming that holobionts are not evolutionary individuals, understood as units of selection (Douglas and Werren 2016, 5).

The realization of these difficulties has led some to argue that holobionts should not be considered as biological individuals, but rather as ecological communities wherein the microorganisms that integrate the host's microbiota should be taken as environmental factors for the host's development and functioning (e.g. Moran and Sloan 2015; Douglas and Werren 2016; Hester et al. 2016; Skillings 2016).

Since the holobiont concept became popularized, holobionts have been interpreted in different, sometimes contradictory, ways, with no clear concept shared among different researchers (Booth 2014; Moran and Sloan 2015;

Douglas and Werren 2016; Skillings 2016; O'Malley 2017). The different theses about the individuality of holobionts have generated an agitated debate: are holobionts biological individuals or mere aggregates of independent individuals that interact in an ecological community? Although some argue that saying that holobionts are ecological units already entails recognizing that they are biological individuals, at least in an ecological sense, as ecosystems could be considered individuals (van Baalen and Huneman 2014), this is not the general way the advocates of the ecological-community view characterize holobionts. Thus, I will mainly consider here the view according to which holobionts are not biological individuals. In this approach, for something to be considered a biological individual, it requires a higher degree of cohesion than the one ecological communities manifest (Skillings 2016).

Recently, Chiu and Eberl have developed a new philosophical framework for conceiving holobionts, what they call the eco-immunity account. According to their proposal, holobionts are not biological individuals, but *hybrids* between a host and its microbiota, the latest playing the role of a scaffold of the individuality of the host (Chiu and Eberl 2016). Chiu and Eberl's account represents a new way to address the holobionts debate. On the one hand, they approach the problem from a perspective that combines the importance of the immunological system of the host for defining its individuality with the relevance of the ecological environment the host interacts with. Especially, they pay attention to the relevance of the microbiota for the maturation and maintenance of the immunological system. On the other hand, they use the *hybrid framework* to characterize the holobiont, thus partially escaping from the dichotomy between the holobiont hypothesis and the ecological-community view.

The eco-immunity approach has important and unexplored metaphysical implications. By arguing why the microbiota *is not a proper part* of the host, as well as by concluding that the holobiont is a *hybrid,* Chiu and Eberl give credit to the metaphysical dimension of the relation of parthood between the host and its microbiota, an aspect that illustrates the general tendency to incorporate metaphysical considerations regarding mereology and whole-part relations in the philosophical analysis of biological individuality (Lidgard and Nihart 2017, 26-28).

Chiu and Eberl's account also highlights the higher degree of *cohesion* that the holobiont, insofar as it is a hybrid, possesses. Thus, on the one hand, it is due to the rejection that the microbiota is a proper part of the host that the authors are in the position to claim that the holobiont is not a biological individual but a hybrid. On the other hand, by claiming that the holobiont is a hybrid, they are recognizing that the holobiont is an entity with a higher degree of cohesion than a mere ecological community, insofar as hybrids are entities that realize a developmental capacity that cannot be realized by an ecological community. This is the main reason why the eco-immunity approach can be considered as an original and genuine account of the holobiont that acts, in a way, as a mediator between those positions that argue for the biological individuality of the holobiont in any of its possible dimensions (anatomical, physiological, developmental, evolutionary), and those positions that reject it. In other words, the eco-immunity account argues that the holobiont has a higher degree of cohesion than ecological communities without assuming that it is a biological individual.

2. The eco-immunity account of holobionts

Chiu and Eberl conceptualize the holobiont debate from a new, immunological, framework. It is from this frame that they develop an eco-immunity account of the holobiont wherein holobionts are not biological individuals but hybrids between a host macroorganism and its microbiota. According to Chiu and Eberl, although the host cannot persist or be individualized without the microbiota, both the host and its microbiota are not sufficiently integrated for the holobiont to be considered as a biological individual or as a unit of selection.

The eco-immunity framework is grounded on two recent theories on immunity, namely the Discontinuity Theory and the Equilibrium Model. According to *the discontinuity theory,* the immune system of an organism establishes the conditions that determine what constitutes it, i.e. its constituent parts (*criterion of inclusion*), and also the conditions in which the organism is actively maintained in the face of constant external perturbations (*criterion of persistence*) (Pradeu and Vivier 2016). As Chiu and Eberl claim:

> For the Discontinuity Theory, a criterion of immunogenicity is a criterion of inclusion, i.e. the conditions that determine what constitutes an organism. From an immunological point of view, what makes a microorganism a part of an organism is not its taxonomic or functional properties, but whether it is interconnected with host components through the biochemical interactions of the immune system. [...] Immunity also contributes to host individuality by the active maintenance of immune responses that affect the persistence of organisms. [...] In addition to the management of energy flows, immunity is emerging as an important way living systems resist degeneration in the face of continuous external invasion and perpetual injury, cellular death, and tumors (Chiu and Eberl 2016, 822).

According to *the equilibrium model,* immunity should be conceived of in terms of the types of reactions that are triggered as a response to three different kinds of targets that the immune system faces, namely (Eberl 2016) *intracellular signals* -from virus and some bacteria; *small extracellular signals* -from helminths and tissue injuries; and *large extracellular signals* -from fungi and most bacteria. These three types of immunity are differentiated by the location, size and general kingdom of the microorganisms' signals and the host's injury signals (Chiu and Eberl 2016, 824). Furthermore, insofar as all of the three kinds of stimuli are constantly and simultaneously present, the immune system never rests but it is continuously responding to the different stimuli it is exposed to (Chiu and Eberl 2016, 825). However, only one kind of response prevails over the others depending on the kind of infection or injury. In this regard, the kind of response that prevails inhibits the effects of the others such that the immune responses are 'downregulated':

> As a result, even though each immune response has the potential to eliminate a target, it does not always do so. Whether an immune response is "allowed" to eliminate its stimulant depends on its *context*, set by the other arms of immunity (Chiu and Eberl 2016, 825).

The regulated mutual inhibition of the immune system to the stimuli is responsible for the persistence and health of the host. The complete inhibition of one kind of response as well as the constant prevalence of one of them over the others can produce important damages in the organism, compromising its health and persistence. As a consequence, there needs to be a kind of balance among the three classes of responses to stimuli in order for the organism to persist. In Chiu and Eberl's view, the microbiota, together with the immune system of the host, allows the persistence and individuality of the host itself

(Chiu and Eberl 2016, 820). The immunity system allows for the host individuality since it determines the constituents of the host. In this regard, insofar as microorganisms interact with the immune system of the host, they also allow for its individuality and persistence (Chiu and Eberl 2016, 829).

However, despite claiming that immunity allows to establish the individuality of the host by determining its constituents, and although the host immunity is also activated by microorganisms, Chiu and Eberl do not consider microorganisms as constituents of the host. In the eco-immunity framework, microorganisms are *scaffolds* of the host individuality, insofar as they allow for appropriate development of the host (the scaffoldee) by increasing the quality, efficiency, and stability of its developmental processes. Furthermore, microorganisms also favor and sometimes control the maintenance of the immunological system of the host, as well as its proper functioning, including how and when it is activated. In this sense, the microorganisms are *resources* that instantiate and regulate the immunological processes necessary for the host constitution, and they do so in a way that allows its persistence.

However, Chiu and Eberl argue, insofar as most microorganisms are *changeable*, *transient,* and sometimes, *shared* by other systems and processes, they are not continuously interacting with the immune system of the host, and therefore, they are not integrated into a single functioning and reproducing whole. This last claim entails that microorganisms cannot be considered as proper parts of the host, which is the genuine biological individual. According to the authors, since microorganisms are not proper parts of the host, holobionts are not integrated anatomic, metabolic, and immunological organisms, which implies that they are not proper biological individuals.

Holobionts, Chiu and Eberl claim, are *hybrids* resulting from the sum of the scaffoldee and its scaffolds. In this sense, the holobiont is the "heterogeneous, collective entit[y]" (Griesemer 2017, 38) "that realizes a developmental capacity" (Chiu and Eberl 2016, 831). Three main aspects characterize the realization of this developmental capacity. Firstly, the scaffolds play a different developmental role than the scaffoldee. For instance, gut development in mice is possible thanks to the interaction of the host cells with *Bacteroides thetaiotaomicron*, which triggers a proinflammatory response of the host's immune system. As a consequence of this proinflammatory response, the host develops the protective mucus layer of its gut that prevents from infections by other microorganisms (Stappenbeck et al. 2002; Round et al. 2011). Secondly, the scaffolds might be changed by other scaffolds while the developmental capacity still gets realized. For instance, under lab conditions, *B. thetaiotaomicron* can be substituted by a different species and still trigger the same response. Finally, changes in scaffolds sometimes correspond to different developmental stages. Birthing and weaning phases, for instance, are characterized by the relationship with certain bacterial species that disappear in adult stages (Funkhouser and Bordenstein 2013; Chiu and Gilbert 2015). Thus, during the period of breastfeeding lactic bacteria, such as *Lactobacilli*, prevails in the gut of the organism. However, after weaning, these bacteria disappear, giving rise to the expansion of a more diverse microbiota that is found in adults (Chiu and Eberl 2016, 827-828).

By paying attention to how the immunological system of the host is co-constructed, the eco-immunological perspective provides a new approach to the holobiont, offering a new way to argue that the microbiota cannot be considered a proper part of the host, as well as rejecting that holobionts are biological individuals. It is precisely due to the rejection that the microbiota is a proper part of the host that the authors claim that the holobiont is not a biological individual, but a hybrid.

In the next section, I focus on the arguments Chiu and Eberl offer in order to reject that microorganisms are proper parts of the host and claim that holobionts are hybrids. It is important to notice here that the rejection of the thesis of holobionts as hybrids does not entail that holobionts are biological individuals. This chapter does not argue for any position regarding the ontological status of holobionts. Rather, I aim to show that once the metaphysical assumptions and implications of Chiu and Eberl's characterization of parthood are made explicit, their argument to illustrate that holobionts are not individuals cannot be properly grounded.

3. Parthood relations in the eco-immunity account

Mereology is the theory regarding the parthood relations that can be given both between a part and a whole and among the parts of a whole (Simons 1987). It is a classic problem in philosophy that has its roots in the Presocratics and is also currently present in metaphysical debates (Varzi 2016). Regarding composite objects, i.e. objects that are composed of other objects, when and how can we properly consider that those objects are proper parts of the composite object? (Lowe 2002). Within the framework of the debate on the ontological status of holobionts, Chiu and Eberl consider the question of parthood to be a question of glue:

> The Holy Grail in the problem of holobiont individuality has been a "special glue" that holds the macro- and microorganisms together. The absence of such a glue, i.e. the disintegration of the holobiont, seems to indicate that holobionts are not biological individuals (Chiu and Eberl 2016, 829).

In this regard, the question about parthood refers to the question on whether the 'special glue' among the parts of the holobiont (i.e. microbiota and macroorganism) is such that it allows us to consider the holobiont as a composite individual. To address this question, the authors propose what might be called an *immunological criterion of parthood* (henceforth ICP), according to which "what makes a microorganism part of an organism is not its taxonomic or functional properties, but whether it is interconnected with host components through the biochemical interactions of the immune system" (Chiu and Eberl 2016, 822).

For ICP, microorganisms might be proper parts of the macroorganism host in virtue of the *continuity* of host-microorganisms immune interactions. However, Chiu and Eberl argue that, insofar as the immune interactions between the host and its microbiota are not continuous, microorganisms are not proper parts of the host. Therefore, they conclude, holobionts are not composite individuals, but mere hybrids between the host macroorganism (the genuine biological individual) and its microbial scaffolds. Chiu and Eberl offer three arguments to justify their claim. First, the lack of some functional glue that integrates the different components of the host with the components of the microbiota. Second, the fact that the different parts that constitute the microbiota of the host can change and are sometimes transient. And third, that parthood is just a special case of scaffolding. In the following sections, I explore the metaphysical assumptions underlying, as well as the metaphysical implications that follow from each of these arguments. In doing so, I will argue that these arguments are not properly grounded and that they give rise to counterintuitive conclusions regarding proper parts when applied to other parts of the host such as organs.

3.1. Causal-functional integration is a kind of glue

As outlined above, for ICP, the fact that microorganisms might play some functional roles for the hosts they interact with cannot be considered as a criterion of parthood. In this regard, Chiu and Eberl claim that, although there are cases wherein microorganisms are believed to be indispensable and highly integrated in the development, metabolism, adaptation, and reproduction of the host, these cases do not show that microorganisms are proper parts of the host: "these fascinating cases at most dispel the notion that macroorganisms are self-sufficient without microorganisms. They fall short from showing that holobionts are causally integrated metabolic or reproductive wholes" (Chiu and Eberl 2016, 821).

The reason why holobionts are not fully reproductively or metabolically integrated, they argue, is that in most cases of host-microorganisms interactions, microorganisms have *negative* or *no effects* on the host phenotype and on its reproduction (Chiu and Eberl 2016, 821). For Chiu and Eberl, this kind of situations would even include those cases where microorganisms completely substitute the organ of a host (e.g. the "digestive tract" co-developed by the *Riftia tubeworm* and its chemosynthesizing bacteria), thus playing a vital function for it (Chiu and Eberl 2016, 821).

So, Chiu and Eberl conclude, microorganisms are not proper parts of the host insofar as, on the one hand, the causal functional interaction they play with the host does not meet the ICP; and, on the other hand, in the majority of interactions between the microbiota and the host, there are microorganisms that play a negative role or no role at all. But ICP entails problematic ontological implications when applied to other constituents of the host, such as organs, since it might lead us to the counterintuitive conclusion that organs are not proper parts of the host.

At least in principle, it seems a natural assumption that organs are proper parts of macroorganisms. Our lungs, our heart, our liver, and our digestive tract are generally considered as proper parts of us. Furthermore, those organs are individualized on the basis that they perform a particular function for organisms that bear them: the heart pumps blood, the lungs perform gas exchange between the blood and the air, the liver produces bile… Organs, therefore, are proper parts of organisms that can be properly individualized by means of their functions (see Chapter 2, § 4.1). In this regard, if there are cases where microorganisms can substitute a complete organ of a host, one might wonder on which grounds the microorganisms that perform the function of such an organ should not be considered as proper parts of the host, whereas an organ is considered to be a proper part of the host due mainly to the function it performs. Why is not functional integration a parthood criterion for microorganisms in the same way as it is a criterion of parthood for other elements of the host?

Chiu and Eberl might reply that such functional structures constituted by microorganisms include some microorganisms with a negative effect or no effect at all. However, having a negative or no effect at all does not seem to be fundamental when referring to the criterion of parthood as applied to organs. In fact, there are many functional units in the organism that contain proper parts that do not perform its function. For instance, the cysts of an organ are parts of the organ, but they do not perform any function that contributes to the organ’s function. Furthermore, there does not seem to be any problem in accepting that macroorganisms can have proper parts, such as organs, with deleterious functions, or with no function at all. In this respect, few would argue that the appendix, for instance, is not a proper part of us. Yet, it is not clear that the appendix has any function. In the same line, it is also possible to refer to cases of organs that do not perform their function properly, i.e. cases of malfunctions (Millikan 1989; Mossio et al. 2009; Saborido 2013). In diabetes, the pancreas does not produce enough insulin (or no insulin at all) to properly regulate the levels of sugar in an organism, for instance. However, in cases of malfunction, the organ is still considered to be a proper part of the macroorganism.

It seems unclear, in Chiu and Eberl's account, why microorganims deserve a different treatment than other proper parts of the body such as organs. In fact, by rejecting the causal-functional integration of the microorganisms with the host, Chiu and Eberl seem to apply two different criteria of parthood for the different entities that compose the holobiont, namely microorganisms and the zygotically-derived components of the host. However, it is not clear why this has to be so: what is exactly the reason why the criterion of parthood that would apply to a malfunctioning pancreas or an appendix would not equally apply to the persistent elements of the microbiota?

My objection might be counter-argued in a twofold way. Firstly, one might argue that the criterion of parthood applied to the zygotically-derived components of the host is different from that applied to its microbiota due to genetic reasons: all the elements that derive from the zygote (tissues, organs, etc.) share the same genome, and thus the criterion of parthood that applies to them might be different from the criterion of parthood that applies to the members of the microbiota. But this argument does not seem to be convincing. On the one hand, because not all the tissues and organs that derive from the zygote are genetically homogeneous (Dupré 2010, 2012), and still we consider them as proper parts of the host. And, on the other, because in the case of transplants, it seems that one would not have any problem in recognizing that the new organ is a proper part of the individual, even using immunological criteria, despite the fact that it is not genetically derived from the zygote (Pradeu 2012).

The denial of the existence of functional integration between the host and its microbiota might also be justified on the basis of ICP. Recall that, according to ICP, microorganisms are parts of the host only when they are interconnected with the host's components through the *continuous* biochemical interactions of the immune system (§ 2). On these grounds, Chiu and Eberl conclude that ICP shows that microorganisms are not proper parts of the host since their interactions are *discontinuous*, that is, the microorganisms that at time t_1 interact with the immunological system of the host might disappear (or be substituted by microorganisms of different species) at a later time t_2.

Appealing to discontinuity, however, does not seem to be a criterion to reject parthood relations between the microbiota and the host either. In fact, to argue in this direction might put in jeopardy the consideration of other components of the host as being proper parts of it. In this regard, *cellular turnover* is a well-known phenomenon that affects all macrobes by means of which their cell types get constantly renewed and replaced by new cells. Despite the constant changes in cells, there seems to be no reason to claim that cells are not proper parts of the macrobe. The examples Chiu and Eberl provide to illustrate the discontinuity of the microbiota, which are mainly cases where the microorganisms play an active role in the development of the host, do not require a different treatment than the one required by cellular turnover: the replacement of old dead cells by new ones is a process that affects all our organs, and it does not jeopardize the fact that those organs are functionally integrated in our bodies. It is difficult to see why the case of the microbiota should require a different treatment.

3.2. Proper parts of an object can change

Chiu and Eberl might offer a counter-argument to our objection that cellular turnover and what can be called 'microbiota turnover' do not seem to require a different treatment. They might claim that there is still a substantial difference between both processes. In the case of microbiome turnover, certain species that used to interact with the immunological system of the host stop interacting after a certain time, whereas in the case of cellular turnover the same cell types are constantly being produced. In fact, Chiu and Eberl offer different arguments that are based on this *changeable* character of the species of microbiota that interacts with a particular host of a given species, in order to argue that the microbiota is not a proper part of it. In the remainder of this section, I pay attention to these arguments, namely contingency, interchangeability, and sharing, in order to illustrate that, from a metaphysical perspective, they are not properly grounded.

Contingency

Besides the lack of a functional glue between the host and its microbiota, the second argument offered by Chiu and Eberl concerning the rejection of the microbiota as a proper part of the host is that although microorganisms are necessary for the host to perform processes such as development, they are also *contingent*:

> The regulation of immune responses microorganisms is necessary for the maintenance of host individuality, but *specific* microorganisms responsible for the regulation are highly *contingent* (Chiu and Eberl 2016, 825. Emphasis added).

Contingency here refers to the fact that the *species* of microbiota that bring about some of the processes of the host, such as its development, might differ among *different hosts* of the *same species*. In Chiu and Eberl's view, this contingency should make us refrain from conceiving microorganisms as proper parts of the host. Chiu and Eberl are right to observe that the particular species of microorganisms that interact with the hosts of a given species are highly variable and depend on various external conditions, such as environmental pressures, kinds of social relationships of the host, diet, etc. However, the contingency of the microbiota does not seem to entail that it is not a proper part of the host.

Let's consider that the microbiota of species MS_1 interacts with the host H_1 of species HS, while the microbiota of species MS_2 interacts with the host H_2 of the same species HS. The consideration of whether the microbiota of species MS_1 is a proper part of the host H_1 is independent of whether the microbiota of species MS_2 is a proper part of the host H_2. To establish parthood relations, it is not relevant that both hosts belong to the same species, that is, what is a proper part of one host is independent of what is a proper part of a different host. In this case, as long as the microbiota is playing a particular function within the host, it is part of it, regardless of whether there is a different species of microbiota that is playing the same function for a different host of the same species. In this respect, therefore, some additional reasons should be provided to justify the claim that the contingent character of the specific microorganisms that interact with a particular host species excludes them for being proper parts of the host. As I will discuss in the next section, xenotransplants are a good illustration of this point.

The interchangeability of microbiota

Another argument developed by Chiu and Eberl to deny that microbiota is a proper part of the host refers to the *interchangeable* character of the microbiota, namely the fact that the species of microbiota that interact *with a particular host* might vary during the lifetime of the host. Thus, a host H that at time t_1 interacts with the species of microbiota M_1, can, at a later time t_2, interact with a different species of microbiota M_2.

However, there does not seem to be any metaphysical problem in recognizing that proper parts of an object can change at different times at which the object exists, without this implying that those are not proper parts of the object itself. In fact, metaphysical theories of persistence are oriented to explain how proper parts of an object can change while the object remains to be the same. As briefly seen in Chapter 1, two main contemporary metaphysical theories are oriented to explain the phenomenon of change and persistence: endurantism and perdurantism. According to endurantism, entities persist by being ‘wholly present’ i.e. by having all of their parts present at each time at which they exist. In this account, change consists of entities having different spatial parts and/or properties at different times at which they exist (Lowe 1987; Simons 2000). For perdurantism, an entity persists by having different temporal parts each of which exists at different times at which the entity exists. In this account, change means having different temporal parts (with different properties) at different times (Quine 1950; Armstrong 1980; Heller 1990). According to perdurantism, an entity is not wholly present at each time at which it exists. Rather, it ontologically refers to the sum of all its temporal parts. It is a spatially and temporally extended entity such that it has past, present, and future temporal parts.

This metaphysical debate concerning the persistence of objects can also be applied to the case of holobionts. In this regard, let's imagine the holobiont (H+MB), as being a spatially and temporally extended entity that exists from time t_0 to t_{10}.[37] At each time, a particular temporal part of the holobiont is manifested, such that we have ten temporal parts: $TP_1 \ldots, TP_{10}$ (see Figure 8).

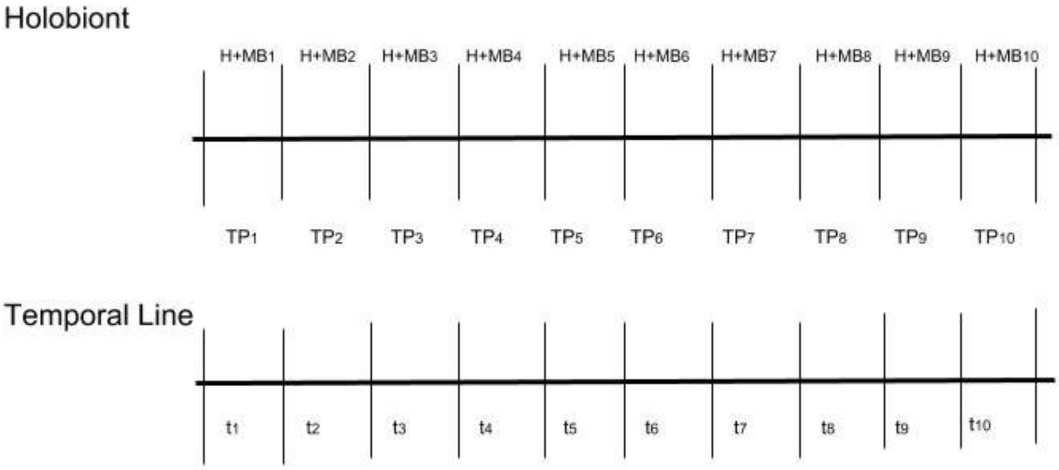

Figure 8. The persistence of the holobiont. This figure represents the holobiont as a spatially and temporally extended entity. In this case, the holobiont persists by perdurance, i.e. by having different temporal parts (TP_1-TP_{10}) that are manifested at different times (t_1-t_{10}) over the interval at which the holobiont exists. Each of the holobiont's temporal part is constituted by a host (H) and its microbiota (MB), which might be different in each temporal part (MB_1-MB_{10}). A change in the holobiont refers to a change in the temporal part that is manifested at a given time. Temporal parts, therefore, are manifested at different times. However, despite not being manifested at a present time, all temporal parts are, in fact, parts of the entity that they constitute.

---

[37] Here, I am using perdurantism to illustrate that the microbiota of a host might vary without this implying that it is not a proper part of the host itself. I am using this metaphysical account of persistence since it allows me to explain my argument in a simpler and better way. However, I consider that there is no problem if the characterization of holobionts and proper parts were made from the framework of endurantism, insofar as I agree with those authors who accept that perdurantists and endurantists accounts of persistence are equivalent (McCall and Lowe 2003, 2006; Miller 2005, 2010).

According to perdurantism, an entity changes insofar as the different temporal parts that constitute it might differ among them, by instantiating, for instance, different qualities. However, this is not problematic. Since each temporal part is manifested at a particular and different time from the others, the object does not possess incompatible properties manifested at the same time, but different properties manifested at different times. In the case of holobionts, we can claim that a particular holobiont is the same entity that exists throughout an extended period of time and, at each time at which it exists, a particular temporal part is manifested. These temporal parts might be different among them by having, for instance, different specific microbiota. Thus, at t_1, the temporal part of the holobiont TP_1 refers to the host H and the microbiota MB_1; at t_2, the temporal part of the holobiont TP_2 refers to the host H and the microbiota MB_2; and so on and so forth. In spite of the microbiota being different at the different times at which the holobiont exists, in all these cases, the microbiota of the holobiont can be considered as a proper part of it, since it is possessed by one or more of the temporal parts that constitute the holobiont.

An illustration of the resource to persistence in order to explain how a biological entity can be composed of different proper parts at different times while remaining the same entity is Reydon's characterization of species in systematic biology (Reydon 2008. See Chapter 1, 4.1). Reydon claims that, in systematic biology, species refer to phylogenetic lineages constituted by relations of common descent. These phylogenetic lineages are seen as ontological entities that persist by having different temporal parts manifested at different times. Each temporal part might be constituted by different individual organisms. However, all organisms that compose the different temporal parts of the species are proper parts of the species itself regardless of whether they exist at a present time *t* or not. This is the reason why, in systematic biology, past organisms of a given phylogenetic lineage are said to be proper parts of the species in the same way as present organisms of the same phylogenetic lineage are said to be proper parts of it (Reydon 2008; Triviño and Cerezo 2015). In other words, the fact that the *specific* organisms of the species might change does not imply that those organisms are not proper parts of the species.

If the metaphysical debate concerning persistence and change is applied to the case of holobionts, it would be possible to claim that, *at least during the period of time* at which the microbiota interacts with the host, the microbiota can still be considered as a proper part of it, especially since they might play essential developmental functions. In this sense, instead of claiming that the microbiota is not a proper part of the host H, insofar as at different times at which H persists, i.g. t_1 and t_2, the microbiota is different, i.e. MB_1 and MB_2, respectively, it might be equally argued that H is composed at different times t_1 and t_2 by different proper parts, that is MB_1 and MB_2, respectively.

Furthermore, and despite the previous argument, it can also be claimed that, although the *species* of microbiota might vary among different hosts of the same species or during the lifetime of a particular host, it is never the case that the host completely eliminates the microorganisms it interacts with. As Chiu and Eberl recognize, the absence of microorganisms is pathological (Chiu and Eberl 2016, 827). Furthermore, the microorganisms that can play the roles of maintenance and development of the host are mainly limited to a few phyla. Not all species of microbiota are able to perform the accurate functions for the hosts (Dethlefsen et al. 2007), something that Chiu and Eberl also recognize:

> Eventually, the host puts strong selection on the microorganisms. Most of the mammal colon bacteria belong only to two phyla, the *Firmicutes* and the *Bacteroidetes*, among the many that exist [...]. Despite the large variation at the species level and the relative stochasticity of microbial colonization when compared to the carefully orchestrated development of the fetus, the types of bacteria and the metabolites they produce, which end up in specific niches of the host, are almost always the same [...] Hence, the structure of the immune system, as well as the structure of the niche (broadly defined as the sum of the selecting parameters, which include tissue, physico-chemical parameters, immunity,...) in the host tissues, select for particular microbial phyla [...] (Chiu and Eberl 2016, 828).

Concerning the limited phyla that can play the roles of maintenance and development of the host, recent empirical evidence strongly supports the idea that "the players might change while the game still remains" (Dolittle and Booth 2017), i.e. that the taxa that compose the microbiota that interacts with a host are less relevant for its maintenance than the specific *functions* that those microorganisms play (Taxis et al. 2015; Lemanceau et al. 2017; Doolittle and Booth 2017; Doolittle and Inkpen 2018). This last empirical evidence can be interpreted, *contra* Chiu and Eberl, as supporting the thesis that the microbiota is a proper part of the host. To see how this might be so, let's consider an analogy concerning organs. Some of our organs might be replaced by structures that play *the same function* but come from different organisms. Transplants are a good example. Our kidneys, for instance, are interchangeable with the kidneys of other humans, provided that they are histocompatible. Furthermore, a lot of studies are taken place in order to make viable the possibility of xenotransplants, wherein the organ that is transplanted to a particular human receptor comes from an animal of other species, especially pigs. If Chiu and Eberl criterion of parthood is employed in these cases, then we need to conclude that transplanted organs are not proper parts of the host.

However, it might seem counterintuitive to deny organs their status as proper parts of the human body on the basis that they are interchangeable by other structures. By analogy, therefore, it does not seem metaphysically grounded to deny microorganisms their status as proper parts of the host on the basis that they are highly interchangeable. What is important is not the species of microorganisms that interact with the host, but that those microorganisms are able to perform the proper functions for it (e.g. Taxis et al. 2015; Hester et al. 2016). Interchangeability, thus, is not a reason to deny an entity its status as a proper part of a bigger whole.

Sharing

Besides contingency and interchangeability, another reason Chiu and Eberl offer to reject microorganisms as proper parts of the host is the fact that, differently to organs, the microbiota is sometimes shared by two (or more) different macroorganisms instead of being possessed by one macrobe exclusively. They illustrate this idea by appealing to the phases of birthing and weaning:

> Microorganisms are often passed on from individuals around us, especially from the mother. [...] Microorganisms are horizontally transferable through frequent social interactions and also in turn influence social behavior (Archie and Tung 2015). These studies might also show that resident microorganisms are shared between collectives instead of exclusively owned by single individuals (Chiu and Eberl 2016, 832-833).

According to Chiu and Eberl, since the microbiome is shared by different hosts in the community, it cannot be considered to be a proper part of the host. It is true, as they remark, that most of the members of the microbiome are not ‘born with us’, or ‘developed with us’, as it happens with our organs, and that this creates a difference between the two. Nonetheless, it is not clear that *being shared* and *being a proper part* should necessarily be mutually exclusive properties. In this regard, metaphysical theories of persistence also take into account cases of overlapping “where A overlaps B just if there is some C such that C is part of A and part of B” (Miller 2010, 44). A case example used to illustrate overlapping is that of the statue and the lump of clay it is made of (Lowe 2002; Miller 2010). Both the statue and the lump of clay can be recognized as two different persistent entities insofar as they are characterized by different properties. Thus, both have different persistence conditions. For instance, a lump of clay can survive events that a statue cannot, such as being squashed flat; and a statue can survive events that a lump of clay cannot, such as having the clay from which it is sculpted replaced by bronze (Miller 2010). Despite having different properties, during the time(s) the statue and the clay coincide they share the same intrinsic properties, such as being brown, for instance. The metaphysical theory of perdurantism has explained cases of overlapping such as this one. According to perdurantism, both the statue and the lump of clay are different entities composed of different temporal parts. However, there are times at which they both share the same intrinsic properties insofar as they share the temporal parts that constitute them at those times.

Applying this metaphysical argument to the case of holobionts, it can be argued that, if it were the case that microorganisms were shared, *at the same* *time*, by two different hosts, such as in the case of Siamese twins that shared different parts or organs, there would not be a metaphysical problem in claiming that these microorganisms are proper parts of the two different hosts at the same time. What would happen, in this case, is that both hosts shared a common temporal part at a particular time, and that temporal part is characterized by having a particular microbiota that is the one that interacts with and characterizes both hosts at the same time.

However, when Chiu and Eberl talk about the possibility for microorganisms to be shared by two different hosts, they are not referring to the idea that microorganisms are possessed *at the same time* by two different hosts. By 'being shared' the authors refer to the idea that the microorganisms that once interacted with a particular host, *pass on to a different host* and interact with it at a different time. In this regard, I consider that it is also possible to claim, without generating problematic metaphysical implications, that the microorganisms that pass from the mum to the offspring during birthing and weaning, for instance, are no more proper parts of the mum, but they are proper parts of the newborn. It is not really relevant, when talking about parthood, that the microorganisms that were part of a host at time t_1 are, at a different time t_2, proper parts of a different host. Again, what is a proper part of a composite object is independent of what currently is, was, or will be, a proper part of a different composite object at a different time.

This kind of metaphysical reasoning concerning the possibility for an object that is a proper part of a composite object at time t_1, to be a proper part of a different composite object at time t_2, also illuminates, in my view, the parthood relations that appear in transplants and transfusions. If someone donates a kidney to her child, for instance, the kidney might be said, without contradiction, to have been a proper part of the parent until t_1 and a proper part of the child at t_2. In fact, as James Griesemer has convincingly argued, all cases of reproduction require a stage of material overlapping between parents and the offspring, i.e. that parts that used to belong to the progenitors in the past do now belong to their offspring (Griesemer 2000, 2014, 2016):

> One clue to a concept of evolutionary minimum relevant sameness is that increase in numbers due to multiplication is always multiplications with "material overlap". [...] Material overlap means that some of the parts of the offspring were once parts of the parents. Material overlap is the relation that results from the flow of matter which creates genealogy (Griesemer 2000, S359).

As I see it, the existence of this material overlap would prove precisely the opposite of what Chiu and Eberl assume it proves, unless the authors are prone to accept that every entity that is generated through reproduction lacks proper parts. Alternatively, they might eventually argue that the criterion of parthood that is applied to the elements of the microbiota should be different from that applied to other elements of the body of the host. But again, they should give a further reason to justify why this is so. Metaphysically speaking, there does not seem to be any reason, *a priori*, to suspect that the microbiota should require a different treatment than any other proper part of an animal or a plant.

3.3. Parthood is not a special case of scaffolding

Finally, another problem I consider in Chiu and Eberl's account of the holobiont is that it is not clear whether the arguments they offer to reject the microbiota to be a proper part of the host refer to resident microbiota, i.e. the elements of the microbiota that reside in the host during its whole life cycle, or to non-resident microbiota. This is a non-trivial aspect that requires more clarification since it makes a difference regarding the kind of microbiota that can be considered a proper part of the host.

At some points in their paper, Chiu and Eberl seem to be talking about resident microorganisms. In fact, the gut bacteria that allow phase transitions and host development in mammals are common examples of resident microorganisms. However, when the authors talk about external microorganisms that might affect the macroorganism, such as microorganisms that stochastically interact with the immune system of the host, they are referring to non-resident, transient microorganisms, whose ontological status in relation to the holobiont is not even agreed upon among the holobiont hypothesis advocates. Thus, some upholders of the holobiont hypothesis have argued that it should exclusively include the permanent microbiota (Lloyd 2017b; Roughgarden et al. 2017), whereas others believe that also transient microorganisms can be included as proper parts of the holobiont (Theis et al. 2016).

Chiu and Eberl seem to rely on this ambiguity, shifting their talk from one type of microorganisms to another, to present their case against considering microorganisms as proper parts of the host. In fact, they base their arguments to claim that resident microorganisms are not proper parts of the host on the difficulty to establish to what extent external microorganisms can be considered as proper parts of it due to their highly contingency and interchangeability. Due to the difficulties in claiming that external microorganisms that interact for short periods of time with the host are proper parts of it, Chiu and Eberl conclude that they are not proper parts of the host. And they extend their conclusion to resident microorganisms as well.

However, the generalization of such a conclusion, insofar as it is based on ambiguity, is not founded. The reason is, again, that the question concerning whether (or not) non-resident microorganisms are proper parts of the host is a different question from the one concerning whether (or not) resident microorganisms are proper parts of it. Yet, Chiu and Eberl might reply that their position is justified on the grounds that, even if there are cases wherein the resident microbiota can be considered as a proper part of the host, parthood is nothing more than a special case of scaffolding, along with the lines of Kim Sterelny's work (Sterelny 2010):

> [O]n one extreme of the landscape, resident microorganisms can be considered proper parts of the host organism: when it is not dependent on other microorganisms, is specialized and necessary for the host and the host only. However, the conceptual map shows that 'parthood' is merely a special case of scaffolding (Sterelny 2010) (Chiu and Eberl 2016, 833).

The identification between parthood and scaffolding, notwithstanding, gives rise to further metaphysical problems concerning how parthood is conceived, as well as about the nature of composite objects. Sterelny's paper, to which Chiu and Eberl refer, focuses on the role that environmental resources play in human cognition. In this regard, Sterelny appeals to two different models that explain this role, namely the Clark-Chalmers extended-mind model and the niche construction model. According to *the Clark-Chalmers model*, the environmental resources that play a role in enhancing and amplifying human capacities (e.g. a notebook enhancing memory) *are proper parts* of human cognitive systems. *The niche construction model*, conversely, states that human cognitive capacities depend on, and have been transformed by, environmental resources that *are not proper parts* of human cognitive systems, but scaffolds of them, i.e. external resources given in the environment that enhance those capacities.

According to Sterelny, the Clark-Chalmers model poses some problems insofar as it seems to imply that *all* environmental factors that enhance human capacities are proper parts of human systems. Thus, if glasses enhance human vision, they should be considered as a proper part of the human visual system (Sterelny 2010, 472). However, the niche construction model has weaker ontological commitments: it does not require us to accept that those external resources are proper parts of human systems, but only resources that *help* in the construction of the human cognitive systems:

> The scaffolded mind hypothesis proposes that human cognitive capacities both depend on and have been transformed by environmental resources. Often these resources have been preserved, built or modified precisely because they enhance cognitive capacity. The extended mind hypothesis proposed that human cognitive systems include external components. These components are coupled with human bodies, but not located within human bodies (Sterelny 2010, 472).

Within the niche construction framework, when Sterelny claims that extended minds are limiting cases of environmental scaffolding, he is not referring to Chiu and Eberl's idea that "'parthood' is merely a special case of scaffolding". In fact, Sterelny does not directly talk about proper parts. He highlights the usefulness of the niche construction model insofar as it makes sense of the idea that environmental scaffolding is a *general phenomenon* that applies both to cognitive and non-cognitive capacities. In this regard, Sterelny offers the modification of food resources and the way we eat as examples of non-cognitive capacities that are environmentally scaffolded. However, he argues, no one would talk about an 'extended digestion': a pot is not a proper part of my digestive system, although it enhances digestion by making food to be soft and easily digestive (Sterelny 2010, 468).

In the same line, Sterelny argues that the resources used for enhancing cognitive capacities are scaffolds in the same way as the ones used to enhance non-cognitive capacities. Thus, the extended mind only refers to a case of a more *general phenomenon*, i.e. that of adaptive phenotypic plasticity based on environmental resources. Regarding the enhancement of cognition, there could be cases in which the environmental resources are so individualized and trusted that they could seem to be proper parts of an extended mind. However, Sterelny rejects this idea on the basis of two arguments. On the one hand, these cases are really rare. On the other hand, even if these cases are given, they are still cases of scaffolding, although, maybe, very special ones.

Sterelny's arguments concerning the advantages of the scaffolded mind over the extended mind also illustrate, at least indirectly, that scaffolds are not proper parts of organisms (Sterelny 2010, 473). However, Chiu and Eberl's interpretation of Sterelny's account is different. They seem to be considering that, for Sterelny, all external resources that enhance organismic capacities are scaffolds of organisms, yet, some of these scaffolds, i.e. the more individualized and trusted ones, are also proper parts of them. By means of this interpretation of Sterelny's account, Chiu and Eberl advocate the following idea: start considering all the microbiota as being scaffolds and then argue that some of those scaffolds are *also* proper parts of the individual, whereas others are not (Chiu and Eberl 2016, 833). Chiu and Eberl interpretation of Sterelny's scaffolding is, notwithstanding, problematic. It erases the distinction between parthood and scaffolds, and thus gives rise to conclusions that seem metaphysically difficult to accept. If parthood is a special case of scaffolding, then all proper parts of a composite object are also scaffolds of it, whereas not all of its scaffolds are proper parts of the object.

There are two problems regarding this characterization of scaffolds. Firstly, there is the question of how to determine which elements of a holobiont are scaffolds of the host and which ones are not. In other words, which is the criterion to consider some entities as components of the host and others as scaffolds? Is it genetic, so that those cells that do not share the same genetic makeup in a multicellular organism are scaffolds and not proper parts of the host? Should we then exclude mitochondria as proper parts of the eukaryotic cell? In this regard, it seems that Chiu and Eberl only provide an intuitive idea of what counts as a scaffold, but do not provide any substantive reason to justify why this has to be so. But let's assume we have a criterion to recognize scaffolds. How can we differentiate between scaffolds that are proper parts and scaffolds that are not proper parts? Is it because the scaffolds that are proper parts are more integrated? If this is so, then it seems that the authors need to recognize that there is, in fact, some kind of glue between the microbiota and the host. However, if functional integration is not a criterion for talking of scaffolds that are proper parts, which seems to be what Chiu and Eberl suggest, then they still need to specify the reasons why one should consider that some scaffolds are proper parts, whereas others are not.

4. Conclusion

In this chapter, I have explored the debate on the ontological status of holobionts from a metaphysical perspective. I have argued that, in this debate, one can identify different metaphysical presuppositions and implications that are not directly recognized or addressed by the participants in it. In order to make explicit some of these metaphysical presuppositions and implications, I have focused on Chiu and Eberl's eco-immunity account. As I see it, the eco-immunity approach has important and unexplored metaphysical implications: by arguing why microbiota *is not a proper part* of the host, as well as by characterizing the holobiont as a *hybrid,* Chiu and Eberl discuss metaphysical aspects regarding holobionts that are not so clearly present in those proposals that conceive of the debate as a dichotomy between biological individuals and ecological communities. In particular, their approach gives credit to the metaphysical dimension of the relation of parthood between the host and its microbiota.

By exploring the individuality of the holobiont on the basis of relations of parthood between the host and the microbiota, Chiu and Eberl implicitly approach the debate concerning biological individuality by incorporating metaphysical considerations regarding the criterion that makes an entity to be a proper part of another, composite, one. My main aim in this chapter has been to make those considerations explicit in order to illustrate that they are not in complete agreement with what metaphysicians claim concerning parthood relations and persistence.

In order to argue that holobionts are not individuals, insofar as the microbiota is not a proper part of the host, Chiu and Eberl offer three different arguments, namely 1) there is a lack of glue between the components of the host and the components of the microbiota; 2) the microorganisms that constitute the microbiota of the host are contingent, interchangeable and shared; and 3), parthood is a special case of scaffolding.

As I have argued, these three arguments are not metaphysically well-founded enough to conclude that the microbiota is not a proper part of the host, and therefore, to claim that holobionts are not individuals. The reasons I have offered can be summarized as follows: 1) it is not well justified why functional integration is not a criterion of parthood for the microbiota when it seems to be so in the case of other components of the host such as organs; 2) the fact that the microbiota of a host can change and be shared by other hosts is not enough to reject that the microbiota is a proper part of the host; and 3) the claim that "'parthood' is merely a special case of scaffolding" (Chiu and Eberl 2016, 833) needs more clarification since the very concept of scaffold is unclear.

Conclusions

I started this dissertation with an emphasis on the idea that metaphysical approaches to conceptual problems in biology are something relatively new. This was not the way philosophers of biology used to proceed due to the negative image of metaphysics postulated by the Modern Synthesis and reinforced by logical positivism. The current tendency in the philosophy of biology to metaphysically approach conceptual issues in biology has called the attention of some philosophers that have started to talk of a Metaphysics of Biology (Guay and Pradeu 2017).

Throughout this dissertation, I have explored the interaction between metaphysics and biology as it occurs in the work of philosophers of biology. I have offered an initial categorization of two ways this interaction takes place as well as the main features that characterize it (Chapter 1). In doing so, I have proceeded by addressing two conceptual problems in biology from a metaphysical perspective, namely the classic problem on the characterization of fitness (Chapters 2 and 3), and the more current debate concerning the ontological status of holobionts (Chapter 4).

With regard to fitness, I have offered a characterization of this property using the metaphysical framework of the dispositional theory of causation and appealing to the notion of emergence. In Chapter 2, I have argued that the dispositional theory of causation developed by Mumford and Anjum (2011) offers the conceptual resources (e.g. threshold, interfering and preventing factors, resultant effect, combination of powers) that allow for a better clarification of the ontological status of fitness as a property of individual organisms. And I propose a characterization of fitness as a causal disposition that results from the non-linear combination of environmentally relative functional dispositions oriented towards an effect (i.e. surviving and reproducing) which is reached once the combination of these dispositions exceeds a certain threshold.

The causal dispositional account of fitness, I have argued, offers several advantages over previous metaphysical proposals such as that of fitness as a supervenient property (Rosenberg 1978) or the propensity interpretation of fitness (Mills and Beatty 1979). The reason is twofold. Firstly, it allows to distinguish realized from expected fitness in a more accurate way by appealing to conceptual resources such as that of normal environment, or the distinction between preventing and interfering factors. Secondly, it allows to conceptualize fitness as both a property of organisms that can change throughout their life history and as a context-sensitive property that depends on the environment and on the population in which organisms are embedded.

In Chapter 3, I have further elaborated my metaphysical characterization of fitness as a causal disposition. I have argued that fitness is a strong ontological emergent property of organisms insofar as its causal power, i.e., making organisms able to survive and reproduce in a particular environment and population, cannot be reduced to the causal power of the functional dispositions that constitute its basis. My argument for this conclusion has been twofold. Firstly, the relation between fitness and the functional dispositions that compose it is a causal one: fitness is the effect of the combination of the organism’s functional dispositions, and effects are different from their causes. Secondly, fitness introduces a new causal power into the world, one that can be downwardly exerted.

In Chapter 4, I have approached the debate on the ontological status of holobionts with special attention to the way metaphysical concepts are used in this debate. In particular, I have focused on the eco-immunity account recently developed by Chiu and Eberl (2016), according to which holobionts are not biological individuals nor ecological communities but hybrids between a host and its microbiota. In doing so, I have looked at the metaphysical implications that follow from the notion of parthood on which the eco-immunity account is based. In this regard, I have argued that the characterization of holobionts as hybrids is not metaphysically grounded. I have offered two major arguments for this conclusion. Firstly, there seem to be two different notions of parthood playing a role in this account, namely one that is applied to explain the relation between the host and its microbiota, and another one that is used to explain the relation between the host and the rest of its components, such as organs. Secondly, extending the criterion of parthood applied to the microbiota to the rest of the components of the host implies counterintuitive conclusions, such that some organs will not be considered as proper parts of the host.

My work on fitness and holobionts contribute to characterize the interaction between metaphysics and biology. This is so because they are instances of two ways metaphysics and biology interact, as I characterized them in Chapter 1: Metaphysics *for* Biology and Metaphysics *in* Biology.

Metaphysics *for* Biology takes place when philosophers appeal to metaphysical theories and/or concepts in order to elucidate the kind of ontological reality biological concepts refer to. In this way, the conceptual issue at stake in biology is clarified. In Chapters 2 and 3 I have proceeded in this way when considering the conceptual problem of fitness. I have directly appealed to metaphysics by means of the dispositional theory of causation and the notion of emergence, which I have exposed throughout these chapters, and I have argued that the problems traditionally attributed to fitness, such as the tautological, the mismatch, and the lack of criterion problems, are better solved when fitness is conceptualized from this metaphysical framework and characterized as a strong emergent causal disposition of organisms.

Metaphysics *in* Biology is the kind of interaction that takes place when philosophers of biology pay attention to biological *theories*, *phenomena,* and *practices*, in order to explore the metaphysical consequences that might follow from them. And, in doing so, the philosopher of biology helps to shed light on the biological theories, phenomena, and practices themselves. Metaphysics, in this case, is used in an indirect way, just to make explicit the presuppositions and implications that follow from the particular biological theory that is being considered. In my chapter on the ontological status of holobionts I have proceeded in this way. The role of metaphysics, in this case, has consisted in exploring the metaphysical implications given in the eco-immunity account due to the notion of parthood it implies. In this regard, I have claimed that (1) functional integration can be considered as a criterion of parthood; (2) proper parts of an object can be contingent, changeable, and shared; and (3) parthood is not a special case of scaffolding. In making explicit these metaphysical commitments, I hope to have contributed to the debate on the ontological status of holobionts. At least, it can be seen that the reasons for arguing that holobionts are hybrids are not metaphysically well grounded and, in this sense, the characterization of holobionts as being either individuals or ecological communities should not be discarded.

Metaphysics *for* Biology and Metaphysics *in* Biology are two different ways the interaction between metaphysics and biology takes place. Some features characterize this interaction: firstly, the form of metaphysics that plays a role in it is not the a priori nor the naturalized one, but one that I have labeled as *applied metaphysics*. Applied metaphysics does not a priori delimit the different ways how the world might ontologically be (as per a priori metaphysics), nor does it follow the scientific constraints to offer an ontological worldview (as per naturalized metaphysics). Yet, it considers both the results in metaphysics and science in order to offer a richer ontological characterization of the world. Secondly, there might be cases of *cross-fertilization* where biology can also contribute to improve the metaphysical theories and concepts initially considered by the philosopher. This is the case in fitness, where some biological phenomena associated with it has led me to reconsider some aspects of the dispositional theory of causation. Thus, I have argued that the existence of powers that can dispose *both towards and away* from biological goals, the transformation of fitness over time, or the epistemological role of interfering factors, require to reconsider or expand some aspects of the dispositional theory of causation. Finally, the two kinds of interaction between Metaphysics and Biology I have identified might not be so clearly distinguished due to cases of *fuzziness*. Thus, the kind of interaction that is at play in each particular case rather depends on where the philosopher is making the emphasis.

Throughout the development of this dissertation, questions and problems have emerged, and I would like to finish by briefly introducing some of them, which are research lines in which I would like to continue working in the future. Firstly, I would like to offer a better characterization of applied metaphysics, its role with regards to science, and how it differs from a priori and naturalized metaphysics. Currently, some works in the metaphysical literature (e.g. French and McKenzie 2012, 2015; Le Bihard and Barton 2018) are oriented to defend a form of metaphysics that is not the a priori nor the naturalized one. I consider that a better characterization of applied metaphysics could make a contribution in this line.

Secondly, concepts in the philosophy of biology such as that of *autonomy*, *agency*, and *natural teleology* were used in Chapter 3 to argue for the emergent nature of fitness, since they were crucial when I justified its downward causal power. These concepts are rich and debated ones in philosophy of biology, but since my approach was metaphysical, I just assumed them, without further development. A nice future research line is to study these notions in more detail and explore a reinforcement of my argument through such analysis.

Finally, concerning the current debate on the ontological status of holobionts, I would like to focus in more detail on the characterization of holobionts as individuals, and in particular, to conceive them as emergent individuals that are characterized by emergent properties. I have already developed a preliminary proposal in this line with Javier Suárez.

***

Metaphysics of biology is a promising new approach within the philosophy of biology that has just started to call the attention of philosophers and to be recognized as a different and fruitful perspective when approaching conceptual problems in biology. The interaction between metaphysics and biology is currently under development and in the process of taking shape. Thus, more work needs to be done in order to conceptualize this tendency. I hope this thesis has shed some light on this direction.

References

Abrams, M. (2006). Fitness and Propensity's Annulment? *Biology and Philosophy*, 22: 11–30.

Abrams, M. (2007). What Determines Biological Fitness? The Problem of the Reference Environment. *Synthese*, 166: 21– 40.

Agassi, J. (1975). *Science in Flux*. Dordrecht and Boston: D. Reidel Publishing Company.

Agassi, J. (1996). The Place of Metaphysics in the Historiography of Science. *Foundations of Physics*, 26(4): 483-499.

Alexander, S. (1920). *Space, time and deity*. London: McMillan.

Allen, S. (2012). What matters in (naturalized) metaphysics? *Essays in Philosophy*, 13: 211-241.

Andersson, M. (1982). Female Choice Selects for Extreme Tail Length in a Widowbird. *Nature,* 299: 818-20.

Anjum, L. and Mumford, S. (2018). Dispositions: A Dynamic Theory of Causation. In: Nicholson, D. and Dupré, J. (eds.), *Everything Flows: Towards a Processual Philosophy of Biology*. Oxford: Oxford University Press, pp. 61-75.

Ariew, A. and Lewontin, R. (2004). The Confusions of Fitness. *The British Journal for the Philosophy of Science*, 55(2): 347-363.

Armstrong, D. (1969). Dispositions are Causes. *Analysis*, 30: 23-26.

Armstrong, D. (1980). Identity through time. In: van Inwagen, P. (ed.), *Time and Cause: Essays Presented to Richard Taylor*. Dordrecht: D. Reidel, pp. 67-78.

Armstrong, D. (1983). *What is a Law of Nature?* Cambridge: Cambridge University Press.

Armstrong, D. (1997). *A World of States of Affairs*. Cambridge: Cambridge University Press.

Austin, Ch. (2016a). Is dispositional causation just mutual manifestation? *Ratio*, 29(3): 235-248.

Austin, Ch. (2016b). The ontology of organisms: mechanistic modules or patterned processes? *Biology and Philosophy*, 31: 639-662.

Austin, Ch. (2017). Evo-devo: a science of dispositions. *European Journal for Philosophy of Science,* 7(2): 373-389.

Austin, Ch. and Nuño de la Rosa, L. (2018). Dispositional Properties in Evo-Devo. In: Nuño de la Rosa, L. and Müller, G. (eds.), *Evolutionary developmental biology: a reference guide*. Springer, pp.1-13.

Baetu, T. (2012). Emergence, therefore antireductionism? A critique of emergent antireductionism. *Biology and Philosophy*, 27: 433-448.

Bapteste, E. and Dupré, J. (2013). Towards a processual microbial ontology. *Biology and Philosophy*, 28: 379–404.

Barandiaran, X. and Egbert, M. D. (2014). Norm-establishing and norm-following in autonomous agency. *Artificial Life Journal*, 20(1): 5-28.

Barandiaran, X. and Moreno, A. (2008). Adaptivity: from metabolism to behavior. *Adaptive Behavior,* 16(5): 325-344.

Barker, J. (2009). Defining fitness in natural and domesticated populations. In: van der Werf, J., Graser, H.U., Frankham, R., and Gondro, C. (eds.), *Adaptation and Fitness in Animal Populations*. Heidelberger: Springer-Verlag, pp. 3-14.

Barnes, E. (2012). Emergence and Fundamentality. *Mind*, 121: 873-901.

Baysan, U. and Wilson, J. (2017). Must strong emergence collapse? *Philosophica,* 91: 49-104.

Beatty, J. (1984). Chance and Natural Selection. *Philosophy of Science*, 51(2): 183–211.

Beatty, J. and Finsen, S. (1989). Rethinking the Propensity Interpretation: A Peek Inside Pandora's Box 1. In: Ruse, M. (ed.), *What the Philosophy of Biology Is*. Springer, pp. 17–30.

Beckermann, A. (1992). Supervenience, Emergence, and Reduction. In: Beckermann, A., Flohr, H., and Kim, J. (eds.), *Emergence or Reduction?* Berlín: Walter de Gruyter, pp. 94-119.

Bedau, M. (1997). Weak Emergence. *Philosophical Perspectives*, 11: 375-399.

Bennett, K. (2011). Construction area (no hard hat required). *Philosophical Studies*, 154: 79–104.

Bertolaso, M. and Cerezo, M. (2017). Robustez Biológica. In: Vanney, C., Silva, I., and Franck, J. (eds.), *Diccionario Interdisciplinar Austral*. URL=http://dia.austral.edu.ar/Robustez_biológica.

Bigelow, R. and Pargetter, R. (1987). Functions. *Journal of Philosophy*, 84: 181–197.

Bird, A. (1998a). Dispositions and Antidotes. *The Philosophical Quarterly,* 48: 227–234.

Bird, A. (1998b). *Philosophy of science*. London: University College London Press.

Bird, A. (2007). *Nature's Metaphysics: Laws and Properties*. Oxford: Oxford University Press.

Bird, A. (2011). Philosophy of science and epistemology. In: French, S. and Saatis, J. (eds). *Continuum Companion to the Philosophy of Science*. London: Continuum International Publishing Group, pp, 15-32.

Bird, A. (forthcoming). A Dispositional Account of Causation, with Some Remarks on the Ontology of Dispositions. In: Meincke, A.S. (ed.), *Dispositionalism: Perspectives from Metaphysics and the Philosophy of Science.*

Brandon, R. (1978). Adaptation and Evolutionary Biology. Reprinted from Brandon, R. *Studies in the History and Philosophy of Science, Vol. 9: Adaptation and Evolutionary Theory,* pp.181–206.

Brandon, R. (1990). *Adaptation and Environment*. Princeton: Princeton University Press.

Brandon, R. (1996). Does biology have laws? The experimental evidence. *Philosophy of Science*, 64: S444-S457.

Brandon, R. (2010). Adaptation and Evolutionary Biology. In: Rosenberg, A. and Arp, R. (eds.), *Philosophy of Biology: An Anthology*. Wiley-Blackwell Press, pp. 102-120.

Brandon, R. (2005). The Difference between Selection and Drift: A Reply to Millstein. *Biology and Philosophy*, 20: 153–170.

Brandon, R. and Beatty, J. (1984). The Propensity Interpretation of "Fitness" - No Interpretation is no substitute. *Philosophy of Science*, 51: 342–347.

Brandon, R. and Carson, S. (1996). The Indeterministic Character of Evolutionary Theory: No "No Hidden Variables Proof" but No Room for Determinism Either. *Philosophy of Science,* 63(3): 315–37.

Brown, C., Garwood, M., and Williamson, J. (2012). It pays to cheat: tactical deception in a cephalopod social signaling system. *Biology Letters*, 8: 729-732.

Boogerd, F.C., Bruggeman, F.J., Richardson, R.C., Stephan, A., and Westerhoff, H.V. (2005). Emergence and its place in nature: a case study of biochemical networks. *Synthese*, 145: 131-164.

Boogerd, F., Bruggeman, F., Hofmeyr, J., and Westerhoff, H. (eds.) (2007). *Systems Biology*: *Philosophical Foundations*. Amsterdam: Elsevier.

Boorse, Ch. (1976). Wright on functions. *Philosophical Review,* 85: 70-86.

Boorse, Ch. (2002). A Rebuttal on Functions. In: Ariew, A., Cummins, R., and Perlman, M. (eds), *Functions: New Essays in the Philosophy of Psychology and Biology*. Oxford: Oxford University Press, pp. 63–112.

Booth, A. (2014). Symbiosis, selection and individuality. *Biology and Philosophy*, 29: 657-673.

Bordenstein, S. and Theis, K. (2015). Host biology in the light of the microbiome: Ten principles of holobionts and hologenomes. *PLOS Biology*. DOI:10.1371/journal.pbio.1002226.

Bouchard, F. (2008). Causal processes, fitness and the differential persistence of lineages. *Philosophy of Science*, 75: 560–570.

Bouchard, F. (2011). Darwinism without populations: a more inclusive understanding of the 'survival of the fittest'. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 42: 106–114.

Bouchard, F. and Rosenberg, A. (2004). Fitness, Probability and the Principles of Natural Selection. *British Journal of Philosophy of Science*, 55(4): 693–712.

Boulter, S. (2013). *Metaphysics from a biological point of view*. U.K: Palgrave Macmillan.

Boyd, R. (1999a). Homeostasis, species and higher taxa. In: Wilson, R. (ed.), *Species: new interdisciplinary studies*. Cambridge: MIT Press, pp. 141–185.

Boyd, R. (1999b). Kinds, complexity and multiple-realization: comments on Millikan's historical kinds and the special sciences. *Philosophical Studies,* 95: 67-98.

Brigandt, I. (2015). Evolutionary Developmental Biology and the Limits of Philosophical Accounts of Mechanistic Explanation. In Braillard, P.A. and Malaterre, C. (eds.), *Explanation in Biology: An Enquiry into the Diversity of Explanatory Patterns in the Life Sciences*. Springer, pp. 135–173.

Broad, C.D. (1925). *The mind and its place in nature*. London: Routledge and Kegan Paul.

Burian, R. (1983). Adaptation. In: Grene, M. (ed.), *Dimensions of Darwinism*. Cambridge: Cambridge University Press, pp. 287-314.

Byron, J. (2007). Whence philosophy of biology? *British Journal for the Philosophy of Science,* 58: 409-427.

Callebaut, W. (1993). *Taking the naturalistic turn, or how real philosophy of science is done*. Chicago: University of Chicago Press.

Callebaut, W. (2005). Again, what the philosophy of biology is not. *Acta Biotheoretica,* 53: 93-122.

Callender, C. (2011). Philosophy of science and metaphysics. In: French, S. and Saatsi, J. (eds.), *The Continuum companion to the philosophy of science*. London: Continuum International Publishing Group, pp. 33-54.

Campbell, D. T. (1974). Downward causation in hierarchically organized biological systems. In: Ayala, F. J. and Dobzhansky, T. (eds.), *Studies in the Philosophy of Biology*. Berkeley/Los Ángeles: University of California Press, pp. 179-186.

Carnap, R. (1936). Testability and Meaning. *Philosophy of Science*, 3: 420-471.

Carnap, R. (1959). The Elimination of Metaphysics through Logical Analysis of Language. In: Ayer, A. *Logical Positivism*. New York: Free Press, pp. 60–81.

Cerezo, M. (Manuscript). Pros and Cons of a Causal Dispositionalist Theory of Genes: Parity, Complexity and Simultaneity.

Chakravartty, A. (2007). *A Metaphysics for Scientific Realism: Knowing the Unobservable*. Cambridge: Cambridge University Press.

Chakravartty, A. (2010). Metaphysics between the sciences and philosophies of science. In: Magnus, P. and Busch, J. (eds). *New Waves in Philosophy of Science*. London: Palgrave MacMillan.

Chakravartty, A. (2013a). On the prospects of naturalized metaphysics. In: Ross, D. Ladyman, J. and Kincaid, H. (eds.), *Scientific Metaphysics.* Oxford: Oxford University Press, pp. 27-50.

Chakravartty, A. (2013b). Dispositions for scientific realism. In: Groff, R. and Greco, J. (eds.), *Powers and Capacities in Philosophy: The New Aristotelianism.* Routledge, pp.113-127.

Chakravartty, A. (2017). Saving the scientific phenomena: what powers can and cannot do. In: Jacobs, J. (ed.), *Putting Powers to Work*. Oxford: Oxford University Press, pp. 24-37.

Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.

Chalmers, D. (2006). Strong and Weak Emergence. In: Clayton, P. and Davies, P. (eds.), *The Re-Emergence of Emergence: from Science to Religion*. Oxford: Oxford University Press, pp. 244-256.

Chiu, L. and Eberl, G. (2016). Microorganisms as scaffolds of biological individuality: An eco-immunity account of the holobiont. *Biology and Philosophy*, 31: 819-837.

Chiu, L. and Gilbert, S. (2015). The Birth of the Holobiont: Multi-species Birthing Through Mutual Scaffolding and Niche Construction. *Biosemiotics*, 8(2): 191-210.

Choi, S. (2003). Improving Bird's Antidotes. *Australasian Journal of Philosophy*, 81: 573–580.

Choi, S. (2005). Dispositions and Mimickers. *Philosophical Studies*, 122: 183–188.

Choi, S. (2006). The Simple vs. Reformed Conditional Analysis of Dispositions. *Synthese*, 148: 369–379.

Choi, S. (2008). Dispositional Properties and Counterfactual Conditionals. *Mind*, 117: 795–841.

Choi, S. and Fara, M. (2018). Dispositions. In: Zalta, E. (ed.), *The Stanford Encyclopedia of Philosophy.*
URL:<https://plato.stanford.edu/archives/fall2018/entries/dispositions/>.

Christensen, W. and Bickhard, M. (2002). The process dynamics of normative function. *The Monist*, 85(1): 3-28.

Clarke, E. (2010). The Problem of Biological Individuality. *Biological Theory*, 5: 312–325.

Clarke, E. (2012) Plant Individuality: A Solution to the Demographer's Dilemma. *Biology and Philosophy*, 27(3): 321–361.

Clarke, E. (2013). The multiple realizability of biological individuals. *The Journal of Philosophy*, 8: 413-435.

Clayton, P. (2006). Conceptual Foundations of Emergence Theory. In: Clayton, P. and Davies, P. (eds.), *The Re-Emergence of Emergence: from Science to Religion*. Oxford: Oxford University Press, pp. 1-31.

Collier, J. (1988). Supervenience and reduction in biological hierarchies. *Canadian Journal of Philosophy*, 18: 209-234.

Collier, J. and Muller, S. (1998). The Dynamical Basis of Emergence in Natural Hierarchies. In: Farre, G. and Oksala, T. (eds.), *Emergence, Complexity, Hierarchy and Organization Selected and Edited Papers from the ECHO III Conference*.

Conee, E. and Sider, Th. (2005). *Riddles of Existence: A Guided Tour of Metaphysics*. Oxford: Oxford University Press.

Creath, R. (2014). Logical empiricism. In: Zalta, E. (ed.), *The Stanford Encyclopedia of Philosophy.*
URL: https://plato.stanford.edu/archives/spr2014/entries/logical-empiricism/.

Cummins, R. (1975). Functional Analysis. *Journal of Philosophy*, 72: 741–765.

Cummins, R. (2002). Neo-teleology. In: Ariew, A., Cummins, R., and Perlman, M. (eds.), *Functions.* Oxford: Oxford University Press, pp. 157-172.

Depaul, M. (1987). Supervenience and moral dependence. *Philosophical Studies*, 51(3): 425-493.

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. 1st edition. London: John Murray.

Davidson, D. (1980 [1970]). Mental Events. In: Block, N. (ed.), *Readings in Philosophy of Psychology*. Cambridge: Cambridge University Press, pp. 107-119.

Davies, P. (1994). Troubles for direct proper function. *Noûs*, 28: 363-381.

Dethlefsen, L., McFall-Ngai, M., and Relman, D. (2007). An ecological and evolutionary perspective on human–microbe mutualism and disease. *Nature,* 449: 811–818.

Dobzhansky, T. (1935). A critique of the species concept in biology. *Philosophy of Science*, 2: 355-355.

Dobzhansky, T. (1954). A review of some fundamental concepts and problems of population genetics. *Symposia on Quantitative Biology*, 20: 1–15.

Doolittle, W.F. and Booth, A. (2017). It's the song not the singer: an exploration of holobiosis and evolutionary theory. *Biology and Philosophy*, 32: 5-24.

Doolittle, W. and Inkpen, A. (2018). Processes and patterns of interaction as units of selection: An introduction to ITSNTS thinking. *Proceedings of the National Academy of Sciences,* 115(16): 4006-4014.

Dorato, M. and Esfeld, M. (2010). GRW as an Ontology of Dispositions. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics,* 41(1): 41-49.

Douglas, A. and Werren, J. (2016). Holes in the hologenome: why host-microbe symbioses are not holobionts. *mBio,* 7(2): 1-7.

Dupré, J. (1993). *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Harvard University Press.

Dupré, J. (2010). The polygenomic organism. *The Sociological Review,* 58(s1): 19-30.

Dupré, J. (2012). *Processes of Life: Essays in the Philosophy of Biology*. Oxford: Oxford University Press.

Dupré, J. (2015). A process ontology: functions are just fast processes and structures are (relatively) slow processes. *Physiology News,* 100: 33-34.

Dupré, J. and Nicholson, D. (2018). A Manifesto for a Processual Philosophy of Biology. In: Nicholson, D. and Dupré, J. (eds.), *Everything Flows: Towards a Processual Philosophy of Biology*. Oxford: Oxford University Press, pp. 3-45.

Dupré, J. and O'Malley, M. (2009). Varieties of living things: life at the intersection of lineage and metabolism. *Philosophy and Theory in Biology,* 1(e003): 1-11.

Eberl, G. (2016). Immunity by equilibrium. *Nature Reviews Immunology,* 16: 524-532.

Ellis, B. (2010). Causal powers and categorical properties. In: Marmodoro, A. (ed.), *The Metaphysics of Powers: Their Grounding and Their Manifestations*. New York: Routledge, pp. 133-140.

Emmeche, C., Koppe, S., and Stjernfelt, F. (1997). Explaining Emergence: Towards an Ontology of Levels. *Journal for General Philosophy of Science,* 28: 83-119.

Engelhard, K. (2010). Categories in the Ontology of Powers: a Vindication of the Identity Theory of Properties. In: Marmodoro, A. (ed.), *The Metaphysics of Powers: Their Grounding and Their Manifestations*. New York: Routledge, pp. 41-57.

Ereshefsky, M. (1998). Species Pluralism and Anti-Realism. *Philosophy of Science*, 65: 103-120.

Esfeld, M. (2004). Quantum entanglement and a metaphysics of relations. *Studies in History and Philosophy of Modern Physics*, 35: 601-617.

Esfeld, M. (2006). The Impact of Science on Metaphysics and Its Limits. *Abstracta*, 2(2): 86-101.

Esfeld, M. (2007). Metaphysics of science: between metaphysics and science. *Grazer Philosophische Studien,* 74: 199–213.

Esfeld, M. (2013). Ontic structural realism and the interpretation of quantum mechanics. *European Journal for Philosophy of Science,* 3: 19-32.

Esfeld, M. (2018). Metaphysics of science as naturalized metaphysics. In: Barberousse, A., Bonnay, D., and Cozic, M. (eds.), *The Philosophy of Science. A companion*. Oxford: Oxford University Press, pp. 142-170.

Fara, M. (2005). Dispositions and Habituals. *Noûs*, 39:43–82.

Fisher, R. (1930). *The Genetic Theory of Natural Selection*. Oxford: Oxford University Press.

Fodor, J. (1974). Special Sciences (Or: The Disunity of Science as a Working Hypothesis). *Synthese*, 28: 97-115.

French, S. (2014). *The Structure of the World: Metaphysics and Representation*. Oxford: Oxford University Press.

French, S. and Ladyman, J. (2003a). Remodeling structural realism: Quantum physics and the metaphysics of structure. *Synthese*, 136: 31-56.

French, S. and Ladyman, J. (2003b). Between platonism and phenomenalism: Reply to Cao. *Synthese*, 136: 73-78.

French, S. and McKenzie, K. (2012). Thinking Outside the Toolbox: Towards a More Productive Engagement Between Metaphysics and Philosophy of Physics. *European Journal of Analytic Philosophy*, 8(1): 42-59.

French, S. and McKenzie, K. (2015). Rethinking Outside the Toolbox: Reflecting Again on the Relationship Between Philosophy of Science and Metaphysics. In: Bigaj, T. and Wüthrich, Ch. (eds.), *Metaphysics in Contemporary Physics (Poznan Studies in the Philosophy of the Sciences and the Humanities, vol. 104)*, pp. 25-54.

Friedman, M. (1999). *Reconsidering logical positivism*. Cambridge: Cambridge University Press.

Funkhouser, L. and Bordenstein, S. (2013). Mom Knows Best: The Universality of Maternal Microbial Transmission. *PLOS Biology,* 11(8): e1001631. DOI:10.1371/journal.pbio.1001631.

Futuyama, D. (1986). *Evolutionary Biology*. 2nd edition. Sunderland: Sinauer.

Gervais, R. (2015). Introductory essay: Metaphysics and Science: a Fickle Relationship. *Philosophica*, 90: 5-22.

Ghiselin, M. (1971). The individual in the Darwinian revolution. *New Literary History,* 3: 113-134.

Ghiselin, M. (1974). A radical solution to the species problem. *Systematic Zoology,* 23: 536-544.

Ghiselin, M. (1997). *Metaphysics and the origin of species*. Albany: SUNY Press.

Ghiselin, M. (2007). The search for the basis of natural classification. *The Monist,* 23: 536-544.

Gilbert, S. and Sarkar, S. (2000). Complexity: Organicism for the 21st Century. *Developmental Dynamics*, 219: 1-9.

Gilbert, S., Sapp, J., and Tauber, A. (2012). A symbiotic view of life: We have never been individuals. *The Quarterly Review of Biology,* 87(4): 325-341.

Gillespie, J. (1977). Natural selection for variances in offspring numbers -a new evolutionary principle. *American Naturalist*, 111: 1010-1014.

Gillet, C. (2002). Strong Emergence as a Defense of Non-Reductive Physicalism: A Physicalist Metaphysics for 'Downward' Determination. *Principia*, 6: 89-120.

Godfrey-Smith, P. (1993). Functions: Consensus without unity. *Pacific Philosophical Quarterly*, 74: 196-208.

Godfrey-Smith, P. (1994). A Modern History Theory of Functions. *Noûs*, 28: 344–362.

Godfrey-Smith, P. (2000). On the Theoretical Role of 'Genetic Coding'. *Philosophy of Science*, 67: 26–44.

Godfrey-Smith, P. (2008). Information in Biology. In: Hull, D. and Ruse, M. (eds.), *The Cambridge Companion to the Philosophy of Biology*. Cambridge: Cambridge University Press, pp. 103-119.

Godfrey-Smith, P. (2009). *Darwinian populations and natural selection*. Oxford: Oxford University Press.

Godfrey-Smith, P. (2015). Reproduction, symbiosis, and the eukaryotic cell. *Proceedings of the National Academy of Sciences*, 112(33): 10120-10125.

Goldstein, J. (1999). Emergence as a Construct: History and Issues. *Emergence*, 1(1): 49-72.

Goldstein, J. (2004). Emergence Then and Now: Concepts, Criticisms, and Rejoinders: Introduction to Pepper's 'Emergence'. *Emergence: Complexity and Organization*, 6(4): 66-71.

Goodman, N. (1954). *Fact, Fiction and Forecast*. Cambridge: Harvard University Press.

Gordon, J., Knowlton, N., Relman, D., Rohwer, F., and Youle, M. (2013). Superorganisms and Holobionts. *Microbe*, 8(4): 152-153.

Gould, S. (1993). *Eight Little Piggies: Reflections in Natural History.* New York: Norton.

Gould, S. and Vrba, E. (1982). Exaptation: A Missing Term in the Science of Form. *Paleobiology*, 8(1): 4-15.

Graves, L., Horan, L., and Rosenberg, A. (1999). Is Indeterminism the Source of the Statistical Character of Evolutionary Theory. *Philosophy of Science*, 66: 140–57.

Griesemer, J. (2000). Development, culture and the units of inheritance. *Philosophy of Science,* 67: S348-368.

Griesemer, J. (2014). Reproduction and the scaffolded development of hybrids. In: Caporael, L., Griesemer, J., and Wimsatt, W. C. (eds.), *Developing scaffolds in in evolution, culture, and cognition*. Cambridge MA: The MIT Press, pp. 23-55.

Griesemer, J. (2016). Reproduction in complex life cycles: A developmental reaction norms perspective. *Philosophy of science,* 83: 803-815.

Griesemer, J. (2017). Landscapes of developmental collectivity. In: Gissis S., Lamm, E., and Shavit, A. (eds.), *Landscapes of collectivity in the life sciences*. London: The MIT Press, pp. 25-48.

Griffiths, P. (1999). Squaring the circle: natural kinds with historical essences. In: Wilson, R. (ed.), *Species: new interdisciplinary studies*. Cambridge: MIT Press.

Griffiths, P. (2001). Genetic Information: a metaphor in search of a theory. *Philosophy of Science*, 68: 394-412.

Griffiths, P. (2008). Philosophy of biology. In: Zalta, E. (ed.), *The Stanford Encyclopedia of Philosophy.* URL: https://plato.stanford.edu/entries/biology-philosophy/

Guay, A. and Pradeu, T. (2017). Right out of the box: How to situate metaphysics of science in relation to other metaphysical approaches. *Synthese,* 1-20.

Guay, A. and Sartenaer, O. (2016). A new look at emergence. Or when *after* is different. *European Journal for Philosophy of Science*, 6(2): 297-322.

Guilford, T. and Dawkins, M.S. (1991). Receiver psychology and the evolution of animal signals. *Animal Behaviour*, 42(1): 1-14.

Haldane, J.B.S. (1932). *The Causes of Evolution*. New York: Longmans, Green and Co., Ltd.

Hall, B. (1996). Baupläne, phylotypic stages, and constraint: why there are so few types of animals. *Journal of Evolutionary Biology,* 29: 215-61.

Hamlett, G.W.D. (1935). Delayed implantation and discontinuous development in the mammals. *Quarterly Review of Biology*, 10: 432-447.

Heller, M. (1990). *The Ontology of Physical Objects: Four-Dimensional Hunks of Matter*. Cambridge: Cambridge University Press.

Heil, J. (2005). Dispositions. *Synthese,* 144: 343–356.

Heil, J. (2010). Powerful Qualities. In: Marmodoro, A. (ed.), *The Metaphysics of Powers: Their Grounding and Their Manifestations*. New York: Routledge, pp. 59-71.

Hester, E., Barott, K., Nulton, J., Vermeij, M., and Rohwer, F. (2016). Stable and sporadic symbiotic communities of coral and algal holobionts. *The ISME Journal, Multidisciplinary Journal of Microbial Ecology,* 10: 1157-1169.

Hirsch, E. (1993). Review: Peter van Inwagen's Material Beings. *Philosophy and phenomenological research,* 53: 687-691.

Hodge, M. J. S. (1987). Natural Selection as a Causal, Empirical, and Probabilistic Theory. In: Krüger, L. (ed.), *The Probabilistic Revolution*. Cambridge (MA): MIT Press, pp. 233-270.

Horan, B. (1994). The Statistical Character of Evolutionary Theory. *Philosophy of Science*, 61(1): 76–95.

Horgan, T. (1982). Supervenience and Microphysics. *Pacific Philosophical Quarterly*, 63: 29–43.

Horgan, T. (1993). From Supervenience to Superdupervenience: Meeting the Demands of a Material World. *Mind*, 102: 555-586.

Hull, D. (1965a). The effect of essentialism on taxonomy -two thousand years of stasis (I). *British Journal for the Philosophy of Science,* 15: 314-326.

Hull, D. (1969). What philosophy of biology is not. *Synthese,* 20: 157-184.

Hull, D. (1976). Are species really individuals? *Systematic Zoology,* 25: 174-91.

Hull, D. (1978). A matter of individuality. *Philosophy of Science*, 45: 335-360.

Hull, D. (1980). Individuality and selection. *Annual Review of Ecology, Evolution, and Systematics,* 11: 311-332.

Hull, D. (2002). Recent philosophy of biology: a review. *Acta Biotheoretica*, 50: 117–128.

Humphreys, P. (1997). How Properties Emerge. *Philosophy of Science*, 64: 1-17.

Humphreys, P. (2008). Synchronic and Diachronic Emergence. *Minds and Machines*, 18(4): 431-442.

Huneman, P. (2010). Topological explanations and robustness in biological sciences. *Synthese*, 177(2): 213-245.

Hurst, G. (2017). Extended genomes: symbiosis and evolution. *Interface Focus,* 7(20170001): 1-11.

Hüttemann, A. (2007). Causation, Laws and Dispositions. In: Kistler, von M. and Gnassounou, B. (eds.), *Dispositions and Causal Powers*. Ashgate, pp. 207-219.

Hüttemann, A. (2013). A disposition-based-process-theory of causation. In: Mumford, S. and Tugby, M. (eds.), *Metaphysics and Science*. Oxford: Oxford University Press, pp. 101-122.

Huxley, J. (1942/2010). *Evolution: The Modern Synthesis*. MIT Press.

Jackson, F. and Pettit, P. (1990). Program Explanation: A General Perspective. *Analysis*, 50(2): 107-117.

Jesper, K. (2006). The causal exclusion argument. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 131(2): 458-485.

Johnston, M. (1992). How to Speak of the Colors. *Philosophical Studies,* 68: 221–263.

Kant, I. (1790/1987). *Critique of Judgment*. Indianapolis: Hackett Publishing.

Kauffman, S. and Clayton, P. (2006). On emergence, agency and organization. *Biology and Philosophy*, 21: 501-521.

Kim, J. (1982). Psychological Supervenience. *Philosophical Studies*, 41: 51-70.

Kim, J. (1984). Concepts of Supervenience. *Philosophy and Phenomenological Research*, 45: 153-176.

Kim, J. (1987). “Strong” and “Global” Supervenience Revisited. *Philosophy and Phenomenological Research*, 48: 315-326.

Kim, J. (1989). The Myth of Non-reductive Materialism. *Proceedings and Addresses of the American Philosophical Association*, 63(3): 31-47.

Kim, J. (1990). Supervenience as a Philosophical Concept. *Metaphilosophy*, 21: 0026-1068.

Kim, J. (1993). The Nonreductivist’s Troubles with Mental Causation. In: Kim, J., *Supervenience and Mind: selected philosophical essays*, New York: Cambridge University Press, pp. 336-357.

Kim, J. (1999). Making sense of Emergence. *Philosophical Studies*, 95: 3–36.

Kim, J. (2003). Blocking Causal Drainage and Other Maintenance Chores with Mental Causation. *Philosophy and Phenomenological Research*, 67: 128-153.

Kim, J. (2005). Physicalism, or something near enough. Princeton, NJ: Princeton University Press.

Kim, J. (2006). Emergence: Core ideas and Issues. *Synthese*, 151: 547–559.

Kim, J. (2009). Mental Causation. In: McLaughlin, B., Beckermann, A., and Walter, S. (eds.), *The Oxford Handbook of Philosophy of Mind*. Oxford: Oxford University Press, pp. 29-52.

Kitcher, P. (1981). Explanatory unification. *Philosophy of Science*, 48: 507.

Kitcher, P. (1984a). Species. *Philosophy of Science*, 51: 308–333.

Kitcher, P. (1984b). 1953 and all that: A tale of two sciences. *Philosophical Review*, 93: 335–373.

Kitcher, P. (1993). Functions and Design. *Midwest Studies in Philosophy*, 18: 379-397.

Koons, D., Metcalf, C., and Tuljapurkar, S. (2008). Evolution of delayed reproduction in uncertain environments: a life-history perspective. *The American Naturalist,* 172(6): 797-805.

Kunz, W. (2012). *Do species exist?: Principles of taxonomic classification*. Germany: Wiley Blackwell Press.

Ladyman, J. (1998). What is structural realism? *Studies in History and Philosophy of Science*, 29: 409-424.

Ladyman, J. and Ross, D. (2007). *Everything must go*. Oxford: Oxford University Press.

Ladyman, J. and Ross, D. (2013). The World in the Data. In: Ross, D., Ladyman, J., and Kincaid, H. (eds.), *Scientific Metaphysics*. Oxford: Oxford University Press, pp. 108-150.

Laland, K., Matthews, B., and Feldman, M.W. (2016). An introduction to niche construction theory. *Evolutionary Ecology*, 30(2): 191-202.

Lean, O. (2014). Getting the most out of Shannon information. *Biology and Philosophy,* 29: 395-413.

Le Bihan, B. and Barton, A. (2018). Analytic Metaphysics versus Naturalized Metaphysics: The Relevance of Applied Ontology. *Erkenntnis,* 1-17.

Lemanceau, P., Blouin, M., Muller, D., and Moënne-Loccoz, Y. (2017). Let the core microbiota be functional. *TRENDS in Plant Science,* 22(7): 583-595.

Lewes, H. (1875). *Problems of Life and Mind*. London: Trench and Turbner.

Lewis, D. (1973). *Counterfactuals*. Oxford: Blackwell.

Lewis, D. (1983a). Extrinsic Properties. *Philosophical Studies*, 44: 197–200.

Lewis, D. (1983b). New Work for a Theory of Universals. *Australasian Journal of Philosophy*, 61: 343–77.

Lewis, D. (1997). Finkish Dispositions. *The Philosophical Quarterly,* 47: 143–158.

Lewens, T. (2007). Adaptation. In: Hull, D. and Ruse, M. (eds.), *The Cambridge Companion to the Philosophy of Biology.* Cambridge: Cambridge University Press, pp. 1-21.

Lidgard, S. and Nyhart, L. (2017). *Biological individuality: Integrating scientific, historical and philosophical perspectives*. London: The University of Chicago Press.

Lillie, R. S. (1914). The philosophy of biology: Vitalism versus mechanism. *Science*, 40: 840–846.

Lloyd, E. (2017). Holobionts as units of selection: Holobionts as interactors, reproducers, and manifestors of adaptation. In: Gissis, S., Lamm, E., and Shavit, A. (eds.), *Landscapes of collectivity in the life sciences*. London: MIT Press, pp. 351-367.

Lowe, E.J. (1987). Lewis on perdurance versus endurance. *Analysis*, 47(3): 152-154.

Lowe, E.J. (2002). *A survey of metaphysics*. Oxford: Oxford University Press.

Lowe, E.J. (2006). *The Four-Category Ontology: A Metaphysical Foundation for Natural Science*. Clarendon Press.

Lowe, E.J. (2008). Two Notions of Being: Entity and Essence. *Royal Institute of Philosophy Supplements*, 83(62): 23–48.

Lowe, E.J. (2010). On the Individuation of Powers. In Marmodoro, A. (ed.), *The Metaphysics of Powers: Their Grounding and Their Manifestations*. New York: Routledge, pp. 8–26.

Mackie, J. (1977). Dispositions, Grounds and Causes. *Synthese,* 34: 361–370.

Maier, J. (2014). Abilities. In: Zalta, N. (ed.), *The Stanford Encyclopedia of Philosophy*. URL: https://plato.stanford.edu/archives/spr2018/entries/abilities/.

Marchesi, J. and Ravel, J. (2015). The vocabulary of microbiome research: a proposal. *Microbiome*, 3(31). DOI: 10.1186/s40168-015-0094-5.

Margulis, L. and Fester, R. (1991). *Symbiosis as a source of evolutionary innovation: speciation and morphogenesis*. Cambridge: MIT Press.

Marmodoro, A. (ed.). (2010). *The Metaphysics of Powers: Their Grounding and Their Manifestations*. New York: Routledge.

Martin, C.B. (1994). Dispositions and Conditionals. *Philosophical Quarterly*, 44: 1-8.

Martin, C.B. (1997). On the Need for Properties: The Road to Pythagoreanism and Back. *Synthese*, 112: 193-231.

Matthen, M. and Ariew, A. (2002). Two ways of thinking about fitness and natural selection. *Journal of Philosophy*, 99(2): 55–83.

Mayr, E. (1982). *The growth of biological thought*. Cambridge: Cambridge University Press.

Mayr, E. (1988). *Toward a new philosophy of biology: observations of an evolutionist*. Cambridge, MA: Belknap Press of Harvard University Press.

Mayr, E. (2004). *What makes biology unique?: Considerations on the autonomy of a scientific discipline*. Cambridge: Cambridge University Press.

McCall, S. and Lowe, E. J. (2003). 3D/4D equivalence: the twins' paradox and absolute time. *Analysis*, 63(2): 114-123.

McCall, S. and Lowe, E. J. (2006). The 3D/4D controversy: a storm in a teacup. *Noûs*, 40(3): 570-578.

Macdonald, C. and Macdonald, G. (2009). Emergence and Downward Causation. In: Macdonald, C. and Macdonald, G. (eds.), *Emergence in Mind*. Oxford University Press, pp. 139-168.

McFall-Ngai, M., Hadfield, MG., Bosch, TC., Carey, H.V., Domazet-Loso, T., Douglas, A.E., Dubilier, N, Eberl, G., Fukamu, T., Gilber, S.F., Hentschel, U., King, N., Kjelleberg, S., Knoll, A.H., Kremer, N., Mazmanian, S.K., Metcalf, J.L., Nealson, K., Pierce, N.E., Rawls, J.F., Reid, A., Ruby, E.G., Rumpho, M., Sanders, J.G., Tautz, D., Wernegreen, J.J. (2013). Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences*, 110: 3229–3236.

McGowen, M., Erez, O., Romero, R., and Wildman, D. (2014). The evolution of embryo implantation. *The International Journal of Developmental Biology*, 58: 155-161.

McKitrick, J. (2003). A Case for Extrinsic Dispositions. *Australasian Journal of Philosophy*, 81: 155-174.

McKitrick, J. (2010). Manifestations as Effects. In: Marmodoro, A. (ed.), *The Metaphysics of Powers: Their Grounding and Their Manifestations*. New York: Routledge, pp. 71-85.

McLaughlin, B. (1992). The rise and fall of British emergentism. In: Beckermann, A., Flohr, H., and Kim, J. (eds.), *Emergence or reduction? Essays on the prospects of nonreductive physicalism*. Berlin: Walter de Gruyter, pp. 49-93.

McLaughlin, B.P. (1997). Emergence and Supervenience. *Intellectica*, 2(25): 25-43.

McShea D. (2012). Upper-directed systems: a new approach to teleology in biology. *Biology and Philosophy*, 27: 663-684.

Mellor, D. (1974). In Defence of Dispositions. *The Philosophical Review*, 83: 157–181.

Mellor, D. (2000). The Semantic and Ontology of Dispositions. *Mind*, 109: 757-780.

Melnyk, A. (2013). Can Metaphysics Be Naturalized? And If So, How? In: Ross, D., Ladyman, K., and Kincaid, H. (eds.), *Scientific metaphysics*. Oxford: Oxford University Press, pp. 79-95.

Mill, J.S. (1843). *A system of logic*. London: Parker.

Miller, K. (2005). The metaphysical equivalence of three and four dimensionalism. *Erkenntnis*, *62*(1): 91-117.

Miller, K. (2010). *Issues in theoretical diversity: persistence, composition and time*. Dordrecht, The Netherlands: Springer.

Millikan, R. (1984). *Language, thought, and other biological categories.* Cambridge, MA: MIT Press.

Millikan, R. (1989). In defense of proper functions. *Philosophy of Science*, *56*(2): 288-302.

Millikan, R. (1999). Historical Kinds and the "Special Sciences". *Philosophical Studies*, 95: 45–65.

Mills, S. and Beatty, J. (1979). The Propensity Interpretation of Fitness. *Philosophy of Science*, 46: 263–286.

Millstein, R. (2002). Are random drift and natural selection conceptually distinct? *Biology and Philosophy*, 17: 33–53.

Millstein, R. (2005). Selection vs. drift: A response to Brandon's reply. *Biology and Philosophy*, 20: 171–175.

Millstein, R. (2006). Natural Selection as a Population-Level Causal Process. *The British Journal for the Philosophy of Science*, 57(4): 627-653.

Millstein, R. (2009). Populations as Individuals. *Biological Theory*, 4(3): 267-273.

Millstein, R. (2016). Probability in Biology: The Case of Fitness. In: Hájek, A. and Hitchcock, C. (eds.), *The Oxford Handbook of Probability and Philosophy*. Oxford: Oxford University Press, pp. 601-624.

Millstein, R. and Skipper, R. (2007). Population genetics. In: Hull, D. and Ruse, M. (eds.), *The Cambridge Companion to the Philosophy of Biology*. New York, U.S.: Cambridge University Press, pp. 22-43.

Mitchell, S. (2012). Emergence: Logical, Functional and Dynamical. *Synthese*, 185(2): 171-186.

Molnar, G. (1999). Are Dispositions Reducible? *Philosophical Quarterly*, 49: 1-17.

Molnar, G. (2003). *Powers: A Study in Metaphysics*. Oxford: Oxford University Press.

Morales, J.D. (2015). *The emergence of mind in a physical world.* PhD Thesis, Colombia University.

Moran, N. and Sloan, D. (2015). The Hologenome Concept: Helpful or Hollow? *PLOS Biology*, 13(12): e1002311.

Morange, M. (2006). Post-Genomics, between Reduction and Emergence. *Synthese*, 151(3): 355-360.

Moreno, A. and Barandiaran, X. (2004). A Naturalized Account of the Inside-Outside Dichotomy. *Philosophica*, 73: 11-26.

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

Morgan, L. (1923). *Emergent Evolution.* London: William and Norgate.

Mossio, M., Saborido, C., and Moreno, A. (2009). An Organizational Account of Biological Functions. *British Journal of Philosophy of Science*, 60: 813–841.

Mumford, S. (1998). *Dispositions*. Oxford: Oxford University Press.

Mumford, S. (2004). *Laws in Nature*. New York: Routledge.

Mumford, S. and Anjun, R. (2010). A Powerful Theory of Causation. In: Marmodoro, A. (ed.), *The Metaphysics of Powers: Their Groundings and Their Realizations*. London: Routledge, pp. 139–151.

Mumford, S. and Anjum, R. (2011). *Getting Causes from Powers*. Oxford: Oxford University Press.

Mumford, S. and Anjum, R. (2013). *Causation: A Very Short Introduction*. Oxford: Oxford University Press.

Mumford, S. and Tugby, M. (2013). *Metaphysics and science*. Oxford: Oxford University Press.

Nathan, M. (2012). The varieties of molecular explanation. *Philosophy of Science*, 79: 233–254.

Needham, P. (2009). Reduction and emergence: A critique of Kim. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 146(1): 93-116.

Nicholson, D. and Dupré, J. (2018). *Everything Flows*: *Towards a Processual Philosophy of Biology*. Oxford: Oxford University Press.

Nicholson, D. and Gawne, R. (2015). Neither logical empiricism nor vitalism, but organicism: what the philosophy of biology was. *History and Philosophy of the Life Sciences*, 37: 345-381.

Nuño de la Rosa, L. (2013). El problema de la función en evo-devo. *Contrastes, suplemento,* 18:187-199.

Nuño de la Rosa, L. and Villegas, C. (Manuscript). Chances and propensities in evo-devo: variability, evolvability and random variation.

O'Connor, T. (1994). Emergent Properties. *American Philosophical Quarterly*, 31: 91-104.

O'Connor, T. and Wong, H. Y. (2005). The Metaphysics of Emergence. *Noûs,* 39: 658-678.

Odling-Smee, F., Laland, K., and Feldman, M. (2003). *Niche construction: the neglected process in evolution*. Princeton: Princeton University Press.

O'Malley, M. (2017). From endosymbiosis to holobionts: Evaluating a conceptual legacy. *Journal of Theoretical Biology*, 434: 34-41.

O'Malley, M. and Dupré, J. (2007). Size doesn't matter: Towards a more inclusive philosophy of biology. *Biology and Philosophy*, 22: 155–191.

Oppenheim, P. and Putnam, H. (1958). Unity of Science as a Working Hypothesis. *Minnesota Studies in the Philosophy of Science*, 2: 3-36.

Orr, H. A. (2009). Fitness and its role in evolutionary genetics. *Nature Review Genetics*, 10(8): 531-539.

Orr, T. and Zuk, M. (2014). Reproductive delays in mammals: an unexplored avenue for post-copulatory sexual selection. *Biological Reviews of the Cambridge Philosophical Society*, 89(4): 889-912.

Overton, W. (1991). The structure of developmental theory. In: van Geert, O. and Mos, L. (eds.), *Annals of theoretical psychology*. New York: Planum, pp. 191-235.

Papineau, D. (1996). *The philosophy of science*. Oxford: Oxford University Press.

Pence, Ch. and Ramsey, G. (2015). Is organismic fitness at the basis of evolutionary theory? *Philosophy of Science*, *82*(5): 1081-1091.

Pepper, S. (1926). Emergence. *Journal of Philosophy*, 23: 241-245.

Petrie, B. (1987). Global supervenience and reduction. *Philosophy and Phenomenological Research*, 48: 119-130.

Pigliucci, M. (2007). Do we need an extended evolutionary synthesis? *Evolution*, 61(12): 2743-2749.

Pigliucci, M. and Müller, G. (2010). *Evolution: The Extended Synthesis*. MIT Press.

Pradeu, T. (2012). *The limits of the self: immunology and biological identity*. NY: Oxford University Press.

Pradeu, T. (2016). The many faces of biological individuality. *Biology and Philosophy*, 31(6): 761-773.

Pradeu, T. and Vivier, E. (2016). The discontinuity theory of immunity. *Science Immunology,* 1: aag0479-0479.

Prior, E. (1985). *Dispositions*. Aberdeen: Aberdeen University Press.

Prior, E., Pargetter, R., and Jackson, F. (1982). Three Theses about Dispositions. *American Philosophical Quarterly*, 19: 251–257.

Queller, D.C. and Strassman, J.E. (2016). Problems of multispecies organisms: endosymbionts to holobionts. *Biology and Philosophy*, 31: 855-873.

Quine, W. v. O. (1950). Identity, ostension, and hypostasis. In: Quine, W. v. O., *From a Logical Point of View*. London: Harvard University Press.

Quine, W. v. O. (1960). *Word and Object*. Cambridge: MIT Press.

Quine, W. v. O. (1981). *Theories and Things*. Cambridge, MA: Harvard University Press.

Ramsey, G. (2015). The Causal Structure of Evolutionary Theory. *Australasian Journal of Philosophy*, 94(3): 421-434.

Reshef, L., Koren, O., Loya, Y., Zilber-Rosenberg, I., and Rosenberg, E. (2006). The coral probiotic hypothesis. *Environmental Microbiology*, 8: 2068–2073.

Reydon, T. (2008). Species in three and four dimensions. *Synthese*, 164(2): 161-184.

Richards, R. (2010). *The species problem: a philosophical analysis*. Cambridge: Cambridge University Press.

Rieppel, O. (2005). Modules, kinds, and homology. *Journal of Experimental Zoology: Part B Molecular and Developmental Evolution,* 304: 18-27.

Rieppel, O. (2006). “Type” in morphology and phylogeny. *Journal of Morphology,* 267: 528-35.

Robus, O. (2014). Does science license metaphysics? *Philosophy of Science,* 82(5): 845-855.

Rosenberg, A. (1978). The supervenient of biological concepts. *Philosophy of Science,* 45(3): 368-386.

Rosenberg, A. (1983). Fitness. *The Journal of Philosophy*, 80: 457-473.

Rosenberg, A. (1994). *Instrumental biology or the desunity of science.* Chicago: University of Chicago Press.

Rosenberg, A. (1985). *The Structure of Biological Science.* Cambridge: Cambridge University Press.

Rosenberg, A. and Williams, M. (1986). Fitness as Primitive and Propensity. *Philosophy of Science*, 53(3): 412–418.

Rosenberg, E., Koren, O., Reshef, L., Efrony, R., and Zilber-Rosenberg, I. (2007). The role of microorganisms in coral health, disease and evolution. *Nature Reviews: Microbiology,* 5: 355-362.

Rosenberg, E. and Zilber-Rosenberg, I. (2014). *The Hologenome Concept*. London: Springer.

Rosenberg, E. and Zilber-Rosenberg, I. (2016). Microbes drive evolution of animals and plants: the hologenome concept. *mBio,* 7(2): e01395-15. DOI: 10.1128/mBio.01395-15.

Ross, D., Ladyman, K., and Kincaid, H. (2013). *Scientific metaphysics*. Oxford: Oxford University Press.

Roughgarden, J., Gilbert, S., Rosenberg, E., Zilber-Rosenberg, I., and Lloyd, E. (2017). Holobionts as Units of Selection and a Model of their Population Dynamics and Evolution. *Biological Theory*, 13: 44-65.

Round, J. L., Lee, S. M., Li, J., Tran, G., Jabri, B., Chatila, T., Mazmanian, S. (2011). The Toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota. *Science,* 332: 974-977.

Ruse, M. (2000). Booknotes 15.3. *Biology and Philosophy*, 15: 465–473.

Ryle, G. (1949). *The Concept of Mind.* London: Hutchinson.

Saborido, C. (2013). Natural Norms and Biological Functions. *Contrastes. Revista Internacional de Filosofía: Suplemento,* 18: 113-124.

Saborido, C. (2014). New Directions in the Philosophy of Biology: A New Taxonomy of Functions. *New Directions in the Philosophy of Science*, 235-251.

Sartenaer, O. (2013). Neither Metaphysical Dichotomy nor Pure Identity: Clarifying the Emergentist Creed. *Studies in History and Philosophy of Science. Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*. 44(3): 365-373.

Sartenaer, O. (2015). Emergent Evolutionism, Determinism and Unpredictability. *Studies in History and Philosophy of Science. Part A*, 51: 62-68.

Schaffner, K. (1967). Approaches to Reduction. *Philosophy of Science*, 34: 137-147.

Schaffner, K. (1969). The Watson-Crick model and reductionism. *British Journal for the Philosophy of Science*, 20: 325–348.

Scriven, M. (1959). Explanation and Prediction in Evolutionary Theory. *Science*, 130(3374): 477–482.

Seibt, J. (2016). Process philosophy. In: Zalta, E. (ed.), *The Stanford Encyclopedia of Philosophy*.
URL:https://plato.stanford.edu/archives/win2016/entries/processphilosophy/

Shapiro, L. and Sober, E. (2007). Epiphenomenalism -the Do's and the Don'ts. In: Wolters, G. and Machamer, P. (eds.), *Thinking about Causes: From Greek Philosophy to Modern Physics*. Pittsburgh: University of Pittsburgh Press, pp. 235-264.

Shea, N. (2011). Developmental systems theories formulated as a claim about inherited representations. *Philosophy of Science*, 78: 60-82

Shea, N. (2013). Inherited representations are read in development. *British Journal for the Philosophy of Science,* 64: 1-31.

Shoemaker, S. (1980). Causality and Properties. In: van Inwagen, P. (ed), *Time and Cause.* D. Reidel, pp. 109-135.

Shoemaker, S. (1984). *Identity, Cause, and Mind*. Cambridge: Cambridge University Press.

Shoemaker, S. (1990). Qualities and Qualia: What's in the Mind? *Philosophy and Phenomenological Research*, 50: 109-131.

Shoemaker, S. (2001). Realization and Mental Causation. In: Gillett, C. and Loewer, B. (eds.), *The Proceedings of the Twentieth World Congress of Philosophy*. Cambridge University Press. pp. 23-33.

Shropshire, J. and Bordenstein, S. (2016). Speciation by symbiosis: the microbiome and behavior. *mBio,* 7(2): 1-11.

Silberstein, M. and McGeever, J. (1999). The search for ontological emergence. *The Philosophical Quarterly*, 48(125): 182-200.

Simons, P.M. (1987). *Parts. A Study in Ontology*. Oxford: Clarendon Press.

Simons, P.M. (2000). How to exist at a time when you have no temporal parts. *The Monist*, 83: 419-436.

Skillings, D. (2016). Holobionts and the ecology of organisms: Multi-species communities or integrated individuals? *Biology and Philosophy*, 31: 875-892.

Smith, A. D. (1977). Dispositional properties. *Mind*, 86: 439-445.

Sober, E. (1984). Fact, Fiction, and Fitness: A Reply to Rosenberg. *The Journal of Philosophy*, 81: 372-383.

Sober, E. (1993). *The Nature of Selection: Evolutionary Theory in Philosophical Focus*. Cambridge: University of Chicago Press.

Sober, E. (2001). The Two Faces of Fitness. *Thinking about Evolution: Historical, Philosophical and Political Perspectives*, 2: 309–321.

Soto, C. (2015). The current state of the metaphysics of science debate. *Philosophica*, 90: 23-60.

Sperry, R. W. (1969). A modified concept of consciousness. *Psychological Review*, 76(6): 532-536.

Stamos, D. (2003). *The species problem: biological species, ontology and the metaphysics of biology*. U.K: Lexington Books.

Stappenbeck, T. S., Hooper, L. V., and Gordon, J. I. (2002). Developmental regulation of intestinal angiogenesis by indigenous microbes via Paneth cells. *Proceedings of the National Academy of Sciences USA*, 99: 15451-15455.

Stencel, A. and Proszewska, A. (2018). How Research on Microbiomes is Changing Biology: A Discussion on the Concept of the Organism. *Foundations of Science*, 23:606-620.

Stephan, A. (1999). Varieties of Emergentism. *Evolution and Cognition*, 5: 49-59.

Stephan, A. (2002). Emergentism, irreducibility, and downward causation. Grazer Philosophische Studien, 65: 77-93.

Sterelny, K. (2010). Minds: extended or scaffolded? *Phenomenology and the Cognitive Sciences,* 9: 465–481.

Sider, T. (2001). *Four-Dimensionalism: An Ontology of Persistence and Time*. Oxford: Oxford University Press.

Swinburne, R. (1980). Properties, Causation, and Projectability: Reply to Shoemaker. In: Cohen, L. and Hesse, M. (eds.), *Applications of Inductive Logic*. Oxford: Oxford University Press, pp. 313-320.

Tahko, T. (2008). *The necessity of metaphysics*. PhD Thesis, Durham University.

Tahko, T. (2013). Metaphysics as the first philosophy. In: Feser, E. (ed.), *Aristotle on Method and Metaphysics*. Palgrave-Macmillan, pp. 49-67.

Taxis, T., Wolff, S., Gregg, S., Minton, N., Zhang, C., Dai, J., Schnabel, R., Taylor, J., Kerley, M., Pires, J., Lamberson, W. and Conant, G. (2015). The players may change but the game remains: network analyses of ruminal microbiomes suggest taxonomic differences mask functional similarity. *Nucleic Acids Research,* 43(20): 9600-9612.

Theis, K., Dheilly, N., Klassen, J., Brucker, R., Baines, J., Bosch, T., Cryan, J., Gilbert, S., Goodnight, C., Lloyd, E., Sapp, J., Vandenkoornhuyse, P., Zilber-Rosenberg, I., Rosenberg, E., and Bordenstein, S. (2016). Getting the hologenome concept right: an eco- evolutionary framework for hosts and their microbiomes. *mSystems,* 1(2): e00028-16. DOI: 10.1128/mSystems.00028-16.

Thoday, J.M. (1953). Components of fitness. *Symposia of the Society for Experimental Biology (VII)*, 96-113.

Tolman, C. (1991). For a more adequate concept of development with help from Aristotle and Marx. In: van Geert, P. and Mos, L. (eds.), *Annals of theoretical psychology*. New York: Plenum, pp. 349-362.

Triviño, V. and Cerezo, M. (2015). The metaphysical equivalence between 3D and 4D theories of species. *Revista Portuguesa de Filosofía*, 71: 781-806.

Triviño, V. and Nuño de la Rosa, L. (2016). A causal dispositional account of fitness. *History and Philosophy of the Life Sciences*, 38: 1-18.

Uebel, T. (2016). Vienna Circle. In: Zalta, E. (ed.), *The Stanford Encyclopedia of Philosophy*. URL: https://plato.stanford.edu/archives/spr2016/entries/vienna-circle/

Valsiner, J. (1997). Constructing the personal through the cultural: Redundant organization of psychological development. In: Amsel, A. and Renninger, K. (eds.), *Change and development: Issues of theory, method, and application.* Mahwah: Erlbaum, pp. 27-42.

van Baalen, M. and Huneman, P. (2014). Organisms as Ecosystems/Ecosystems as Organisms. *Biological Theory*, 9: 357-360.

van Cleve, J. (1990). Mind-dust or Magic? Panpsychism versus Emergence. *Philosophical Perspectives*, 4: 215-226.

van der Steen, W. J. (1994). New ways to look at fitness. History and Philosophy of the Life Sciences, 16(3): 479-492.

van Gulick, R. (2001). Reduction, emergence and other recent options on the mind/body problem: a philosophical review. *Journal of Consciousness Studies*, 8: 1-34.

van Inwagen, P. (1990). *Material beings*. New York: Cornell University Press.

Varzi, A. (2016). Mereology. In: Zalta, E. (ed.), *The Stanford Encyclopedia of Philosophy*.
URL:https://plato.stanford.edu/cgibin/encyclopedia/archinfo.cgi?entry=mereology

Waddington, C.H. (1996). The Practical Consequences of Metaphysical Beliefs on a Biologist's Work: An Autobiographical Note. In: Waddington, C.H. (ed.), *Towards a Theoretical Biology 2: Sketches*. Edinburgh: Edinburgh University Press, pp. 72-81.

Wagner, G. (1996). Homologues, natural kinds and the evolution of modularity. *American Zoologist*, 36: 36-43.

Walsh, D. (1996). Fitness and Function. *Philosophy of Science*, 47(4): 553–574.

Walsh, D. (2007). The pomp of superfluous causes: the interpretation of evolutionary theory. *Philosophy of Science*, 74(3): 281–303.

Walsh, D. (2010). Not a Sure Thing: Fitness, Probability and Causation. *Philosophy of Science*, 77: 147–171.

Walsh, D. (2013). Mechanism, emergence, and miscibility: the autonomy of evo-devo. In: Huneman, P. (ed.), *Functions: Selection and Mechanisms*. Springer, Dordrecht, pp. 43-65.

Walsh, D., Lewens, T., and Ariew, A. (2002). Trials of life: natural selection and random drift. *Philosophy of Science*, 69: 429-446.

Wartick, J. (2012). Book review: ‘Material Beings’ by Peter van Inwagen. In: Wartick’s web page. Available via DIALOG: https://jwwartick.com/2012/01/02/mb-pvi/

Wartofsky, M. (1967). Metaphysics as Heuristic for Science. In: Cohen, R. and Wartofsky, M. (eds.) *Boston Studies in the Philosophy of Science, Vol. III*. Dordrecht D. Reidel Publishing Company, pp. 123-172.

Waters, K. (1994). Genes made molecular. *Philosophy of Science*, 61: 163-185.

Waters, K. (2017). No general structure. In: Slater, M. and Yudell, Z. (eds). *Metaphysics in the philosophy of science: new essays*. Oxford: Oxford University Press.

Watkins, J. (1975). Metaphysics and the Advancement of Science. *The British Journal for the Philosophy of Science*, 26: 91–121.

Weber, M. (2001). Determinism, realism and probability in evolutionary theory. *Proceedings of the Philosophy of Science Association,* 83: 213-224.

West-Eberhard, M. J. (2005). Developmental plasticity and the origin of species differences. *PNAS,* 1: 6543-6549.

Wiggins, D. (2016). Activity, process, continuant, substance, organism. *Philosophy,* 91(2): 269-280.

Williams, M. (1970). Deducing the consequences of evolution. *Journal of Theoretical Biology*, 29: 343-385.

Williams, M. and Rosenberg, A. (1985). "Fitness" in fact and fiction. *Journal of Philosophy,* 82:738-749.

Wilson, J. (1999). How Superduper does a Physicalists Supervenience Need to be? *The Philosophical Quarterly*, 49: 33-52.

Wilson, J. (2002). Causal Powers, Forces, and Superdupervenience. *Grazer Philosophische Studien*, 63: 53-78.

Wilson, J. (2013). Nonlinearity and metaphysical emergence. In: Mumford, S. and Tugby, M. (eds.), *Metaphysics and Science*. Oxford: Oxford University Press, pp. 201-229.

Wilson, J. (2016). Metaphysical emergence: weak and strong. In: Bigaj, T. and Wüthrich, C. (eds.), *Metaphysics in Contemporary Physics Poznan Studies in the Philosophy of the Sciences and the Humanities*, vol. 104. Amsterdam/New York: Rodopi, pp. 345-402.

Wilson, R. (1999). Realism, essence and kind: resuscitating species essentialism? In: Wilson, R. (ed.), *Species: new interdisciplinary studies*. Cambridge: MIT Press, pp. 187-207.

Wilson, R. (2007). Levels of Selection. In: Matthen, M. and Stevens, C. (eds.), *Handbook of the Philosophy of Science (Volume 3, Philosophy of Biology)*. Elsevier, pp. 155-176.

Wilson, R. and Barker, M. (2013). The Biological Notion of Individual. In: Zalta, E. (ed.) *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/sum2018/entries/biology-individual/>.

Witherington, D. (2011). Taking Emergence Seriously: The Centrality of Circular Causality for Dynamic Systems Approaches to Development. *Human Development*, 54: 66-92.

Wright, L. (1973). Functions. *Philosophical Review,* 82: 139-168.

Wright, L. (1976). *Teleological explanations: an etiological analysis of goals and functions.* Berkeley: University of California Press.

Wright, S. (1931). Evolution in Mendelian Population. *Genetics,* 16: 97-159.

Zilber-Rosenberg, I. and Rosenberg, E. (2008). Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiology Reviews*, 723-73.