An Ontic Account of Explanatory Reduction in Biology

Dissertation

zur Erlangung des Grades einer

Doktorin der Philosophie (Dr. phil.)

der Philosophischen Fakultät
der Universität zu Köln

vorgelegt von

Marie I. Kaiser, M.A.

Köln, den 02. Mai 2012

Contents

Illustrations

Introduction

The central goal of this book is to provide an understanding of an important element of contemporary biological research practice, namely of *explanatory reduction* (i.e. reductive explanation; see Chapter III).

The topic of reduction(ism) was and still is a much discussed issue in general philosophy of science, in philosophy of mind, and in philosophy of biology. Most notably, it belongs to the classical issues that were discussed when philosophy of biology emerged as a distinct discipline in the 1960s and 70s. Hence, the general topic I am concerned with in this book is not new. However, the specific question that I address, the way in which I approach this question, and the answer I give are novel. Most discussions about reduction in philosophy of biology have focused on two issues: on the one hand, on the question of whether *reductionism* or antireductionism is ultimately correct, for instance, whether it is in principle possible to adequately explain each biological phenomenon in molecular or in physical terms. On the other hand, discussions about reductionism centered on a particular understanding of reduction, namely on Ernest Nagel's (1961) formal model of *theory reduction*. In the last third of the 20th century philosophers put a lot of effort into defending or criticizing the application of Nagel's model to the biological sciences (in particular to the relation between Mendelian genetics and molecular biology).

My analysis in this book differs from these classical disputes in both respects: my aim is neither to defend a certain version of reductionism or antireductionism with respect to biology, nor is it to discuss the topic of reduction within the narrow boundaries of Nagel's model of theory reduction.[1] Instead, I focus on answering a question that I think is prior to discussions about explanatory reductionism (see Chapter II, Section 1), namely what does it mean to explain a biological phenomenon in a reductive manner? I offer an extensive and (hopefully) persuasive answer to this question by developing my account of explanatory reduction in biology. This account presents an alternative way of thinking about epistemic reduction in biology, which does not remain within the Nagelian framework that reconstructs reduction as a relation of logical derivation between theories.[2]

In the last two decades several philosophers have responded to the shortcomings of Nagel's model of theory reduction by abandoning the concept of reduction or the focus on reduction altogether (e.g. Craver 2005, 2007a; Mitchell 2003, 2009; Mitchell/Dietrich 2006; Darden 2005; Bechtel/Richardson 2010). In my view, this is not the right strategy. Reductions (more precisely, reductive methods and reductive explanations) were and remain important

---

[1] In so doing I align myself with authors like Wimsatt (1976a, 2006a, 2007), Sarkar (1998, 2005), and Hüttemann/Love (2011), who also seek to analyze epistemic reduction in a non-Nagelian way and without being primarily interested in defending (anti-)reductionism.

[2] Although I break new ground in developing my account of explanatory reduction in biology I also tie my analysis to previous discussions about reduction(ism) in philosophy of biology. However, I only examine those issues that are relevant to my project.

elements of biological practice. This is not to say that reductive research strategies do not have crucial limitations and that the attempt to explain the behavior of some biological systems in a reductive way may not result in serious inadequacies. But despite their shortcomings reductive methods and explanations also have certain merits, which is why they still do play an important role in contemporary biological practice. This is supported by the fact that current biological research papers show a continuous attention to the topic of reduction(ism). In particular, biologists discuss questions such as "Under which conditions is the application of reductive methods a permissible and profitable research strategy?" and "Can a particular biological phenomenon be adequately explained in a reductive manner?" (for further details see Chapter V). Thus, I think we should adhere to the concept of epistemic reduction because it is an important conceptual tool for capturing significant aspects of biological research practice. What we need is a better understanding of what epistemic reduction in biology is, not the removal of the concept of reduction from philosophy of biology altogether.

My aim in this book is to develop such a better understanding of epistemic reduction. Precisely speaking, I am concerned with a specific type of epistemic reduction, namely with explanatory reduction, respectively with reductive explanation (this focus will be justified in Chapter II). The central question that I seek to answer in my analysis is: what are the features of biological explanations that determine their reductive character? In other words, which characteristics are common to all (or to most) reductive explanations in the biological sciences and allow one to clearly distinguish reductive from non-reductive explanations? What is crucial to my account of explanatory reduction is that the answer I give to this question emerges from a critical reconstruction of biological research practice itself. That is, my answer does not reflect a philosophical ideal of reduction. Rather, it captures paradigmatic and important cases of explanatory reduction from contemporary biological practice, and it accounts for the way biologists currently discuss the merits and "limits of reductionism" (Ahn et al. 2006a, 709; Mazzocchi 2008, 10) or call for a move "beyond reductionism" (Gallagher/ Appenzeller 1999, 79). By taking actual biological practice seriously my analysis provides several novel insights into the central characteristics of reductive explanations. That way, it clarifies and specifies what it means to explain a biological phenomenon reductively.

The project of developing an account of explanatory reduction in biology is valuable by itself because the concept of a reductive explanation has not received sufficient philosophical attention so far. Moreover, it is beneficial because it has the potential to advance debates about explanatory reductionism. Any fruitful discussion about the truth of explanatory reductionism must be based on a clear understanding of what explanatory reduction is. Otherwise misunderstandings and people talking at cross purposes are a daily occurrence. What is more, my analysis also yields an understanding of what the merits and limitations of applying reductive strategies and developing reductive explanations in biological practice are. This knowledge, not only about what reductive explanations are but also about the conditions under which they succeed and fail to be adequate, is of great value to discussing the plausibility of different versions of explanatory reductionism.[3]

In what follows I, at first, give a rough overview of the general structure of my book. Afterwards, I reveal the line of my argumentation in more detail. Let us start with the rough overview. *Chapter I* serves to disclose the meta-philosophical assumptions that underlie my analysis of explanatory reduction. This includes explicating (and justifying) the aim of my analysis, the philosophical methodology by which I develop my account, and the criteria of adequacy that I accept. I will characterize my own account as being descriptive and bottom-up but critical, as being as universal as possible and as specific as necessary, as being normative in a certain way but not in another, and as being potentially useful for science. The purpose of *Chapter II* is to introduce the previous debate about reduction(ism) in the philosophy of biology. However, this introduction will not be a mere overview. Rather, I present what I conceive as the most crucial lessons one should learn from this debate. In doing so, I introduce and specify important concepts and distinctions. Moreover, I show the reader the path I will run in the remaining part of this book. That is, I adduce reasons for why I develop an account of explanatory reduction, rather than an account of ontological reduction, methodological reduction, or theory reduction. In *Chapter III* I critically discuss the two perspectives on explanatory reduction that have been proposed in the philosophy of biology so far, namely Rosenberg's thesis that explanatory reduction is a relation between a higher-level and a lower-level explanation of the same phenomenon and Sarkar's, Hüttemann's, and Love's approach to focus on individual reductive explanations. The result of my critical examination will be that Rosenberg's perspective on explanatory reduction in biology has several shortcomings and that, even though Sarkar's, Hüttemann's, and Love's encounters objections, too, it seems to be the more promising path to run. If one decides to analyze epistemic reduction by examining the reductive character of individual explanations, as I do, the question arises how entangled the issue of reduction becomes with the issue of explanation. I will address this question in *Chapter IV*. In particular, there are two questions that need to be answered differently: first, does the question of what determines the reductive character of a biological explanation (the question of reduction) boil down to the question of what characterizes an adequate explanation (the question of explanation), and second, do debates about the truth of explanatory reductionism depend on specific discussions about explanation (and if yes, on which)?

The results of these four chapters constitute the ground on which I can then, in *Chapter V*, develop my own account of explanatory reduction in biology. I will do so in two major steps. In the first part of this chapter I specify four concepts that occupy center stage in my analysis of the features of reductive explanations, namely the concept of a biological

---

[3] At least if one is interested in in-practice kinds of explanatory reductionism (see Chapter II, Section 1).

system, the concept of its context or environment, the concept of parts of a system, and the concept of levels of organization. On the basis of these conceptual clarifications I can then in the second part of this chapter answer the central question of my book, namely what are the characteristics that determine whether a biological explanation is reductive or not. The main result of my analysis will be that reductive explanations in biological science possess three features (two of which are necessary conditions, one of which is only a typical feature that most reductive explanations exhibit): they display a lower-level character, focus on factors that are internal to the system of interest, and describe only the parts of a system in isolation.

Let us go into the details. The aim of *Chapter I* is to reveal the meta-philosophical assumptions that I presuppose in developing my account of explanatory reduction in biology. I start by pointing out that the goal of my analysis is to understand what explanatory reduction in current biological research practice is (Section 1). One of my main theses is that this aim commits you to account for cases of explanatory reduction that actually occur in biological practice (reduction *in practice*). In other words, it commits you to accept descriptive adequacy as an important criterion of adequacy (Section 1.1). I clarify in which sense descriptive accounts in philosophy of science can be called naturalistic (Section 1.2) and why the best way to develop them is to conduct a philosophical analysis in a bottom-up fashion (Section 1.3). I conclude by specifying what the "empirical data (about biological practice)" are that a descriptive account needs to capture (Section 1.4). In Section 2 I approach a possible objection that proponents of descriptive accounts (like me) encounter, namely why should one care about biological practice in the first place? I respond to this objection by distinguishing two kinds of projects in philosophy of science, which both can be accepted as legitimate, namely descriptive projects and normative projects. I further show that there exists a continuum between purely descriptive and purely normative projects and that the projects which are located on the middle ground of this continuum can be distinguished on the grounds that they satisfy three different dimensions of normativity. In Section 3 I pick up on one of these dimensions of normativity and elaborate it. The central question I examine in this section is what it means that descriptive projects in philosophy of science cannot be *purely* descriptive since philosophers must take up a critical stance on the empirical data that scientific practice provides them with. One of my main theses will be that philosophers should aspire to construct a *coherent* account, for instance, of reductive explanation. This demand for coherency is further analyzed in Section 4. Here I address the question of how much and which kind of pluralism we in fact need. After telling apart two different senses of non-universality (Section 4.1) I propose a second criterion of adequacy that can be condensed to the motto "Try to achieve as much generality as you can get and as much specificity as you need" (Section 4.2). Finally, I discuss the objection that this criterion is not pluralistic enough (Section 4.3). In Section 5 I pick up on the issue of normativity again. The central task of this section is to clarify what it means when some philosophers of science claim that their account is jointly descriptive and normative. Several authors link the normativity of their project to the fact that it aims at disclosing certain epistemic norms. In order to clarify the notion of normativity that is involved here I first specify the concept of epistemic norms in science (Section 5.1). Then I point to an important distinction, namely the one between describing epistemic norms that are in fact accepted in science and positing epistemic norms that should apply to science (Section 5.2). I argue that those philosophers who seek to justify which norms should apply to science face a dilemma: either they commit an is-ought fallacy or they fail to specify the grounds on which their independent justifications are made (Section 5.3). I conclude by pointing out in which sense my own analysis of reductive explanations in biology is normative (Section 5.4). In the last section of this chapter (Section 6) I address the popular question of whether a philosophical account should be judged according to its actual relevance to scientific practice.

In *Chapter II* I introduce the previous debate about reduction(ism) in the philosophy of biology, which has lasted more than fifty years now. But rather than giving a mere chronological overview of the course of the debate, I present what I conceive as the most crucial lessons one should learn from this debate. That way, I introduce and specify important concepts and distinctions that turn out to be significant to my subsequent analysis. In particular, I identify four lessons that one should draw from the previous debate. First, before you discuss the correctness of reductionism you should understand what reduction is (Section 1). Second, if you seek to understand what reduction in actual biological practice is, you should focus on epistemic, rather than on ontological issues (Section 2.3). But you should keep in mind that questions about ontological reduction and about the interrelations between epistemic and ontological issues, nevertheless, are relevant. This is why I will put some effort into explicating the notion of ontological reduction that underlies current discussions about reductionism in philosophy of biology (Section 2.1) and in specifying which relations exist between epistemic and ontological issues (Section 2.2). Third, before you discuss reduction(ism) you need to specify which kind of reduction you are talking about – ontological reduction, theory reduction (Section 3.1), methodological reduction (Section 3.2), or explanatory reduction (Section 3.3).[4] Since the concept of methodological reduction has not attracted much philosophical attention so far I consider it in more detail. I point out that in the current debate what it means to pursue research in a reductive manner is left underspecified (Section 3.2.1). Therefore I analyze the notion of a reductive method (Section 3.2.2) and the relation that exists between reductive methods and reductive explanations (Section 3.2.3). I conclude by arguing that reductive explanations are a more promising target of philosophical analysis than reductive methods because they are more constrained and less discipline specific (Section 3.2.4). The fourth and last lesson to learn is that, finally, it is time to move beyond Nagel's inadequate model of theory reduction and to break new ground in thinking about epistemic reduction in biology.

---

[4] The latter three are subtypes of epistemic reduction.

In order to show this, I first clarify what the target of my criticism is (Section 4.1 and 4.2). Then I adduce different reasons for why Nagel's model is inadequate (Section 4.3). My general line of criticism will be that Nagel's account fails to account for paradigmatic and important cases of epistemic reduction that actually occur in biological research practice (most notably, reductive explanations). One central goal of this chapter is to explain (and to justify) why I develop an account of *explanatory* reduction in biology, rather than an account of ontological, methodological, or theory reduction.

In *Chapter III* I critically examine what I regard as the two perspectives on explanatory reduction that have been proposed in the philosophy of biology so far: on the one hand, Rosenberg's perspective, according to which explanatory reduction is a relation between a higher-level and a lower-level explanation of the same phenomenon (Section 1), and on the other hand, the perspective of Sarkar, Hüttemann, and Love, according to which explanatory reductions are individual reductive explanations (Section 2). The aim of this critical discussion is to reveal implausible theses as well as fruitful insights concerning explanatory reduction in biology that can be utilized in developing my own account. I start with reconstructing the major aspects of Rosenberg's defense of explanatory reductionism in biology (Section 1.1). On the basis of this I can then reconstruct the notion of explanatory reduction that Rosenberg presupposes in his argumentation (Section 1.2). The result will be that Rosenberg treats explanatory reduction as a relation between two specific kinds of explanations of the same phenomenon, but that the exact nature of the process of reduction remains obscure. I conclude by pointing out some serious objections that Rosenberg's notion of explanatory reduction encounters (Section 1.3). In Section 2 I analyze different versions of what I regard as the second main perspective on explanatory reduction. What binds authors like Kauffmann, Wimsatt, Sarkar, Hüttemann, and Love together is that each of them focuses on individual reductive explanations and investigates the constraints of their reductive character. I begin by giving a brief overview of the first insights Kauffman and Wimsatt gained into the reductive character of biological explanations (Section 2.1). Afterwards, I consider Sarkar's work on reductive explanations in genetics and molecular biology (Section 2.2). I explicate what it means that Sarkar calls his account substantive (Section 2.2.1) and why he emphasizes its neutrality with respect to different accounts of explanation (Section 2.2.2). Then I discuss the core of his account, namely the three criteria of reductivity he identifies (Section 2.2.3). I conclude by revealing the major shortcomings of his analysis (Section 2.2.4). In Section 2.3 I discuss the most recent work on explanatory reduction in biology, namely Hüttemann's and Love's analysis of three aspects of reductive explanation. After delineating their methodological framework (Section 2.3.1) I scrutinize their three aspects, namely intrinsicality, fundamentality, and temporality (Section 2.3.2), and clarify in which sense these aspects constrain the reductive character of a biological part-whole explanation (Section 2.3.3). Finally, I disclose the difficulties their analysis encounters (Section 2.3.4). I conclude by arguing that the second perspective, which focuses on individual reductive explanations and analyzes the conditions under which explanations succeed and fail to be reductive, seems to be the much more promising strategy to analyze explanatory reduction in biology.

Since my project is to analyze epistemic reduction by examining the reductive character of individual explanations the question arises how entangled the issue of reduction becomes with the issue of explanation. The main purpose of *Chapter IV* is to reveal the interrelations that exist between the issue of explanation and my account of explanatory reduction. At the outset of this chapter I give a brief overview of the central ideas of the different accounts of explanation that have been discussed in the last decades as well as on the objections that have been raised against them (Section 1). In particular, I discuss two major kinds of accounts of explanation: first, the covering-law (CL) model of explanation (Section 1.1), under which I subsume Hempel's deductive-nomological (DN) model, his inductive-statistical (IS) model, and the unificationist account of explanation advocated by Friedman and Kitcher; second, the causal-mechanical (CM) model of explanation (Section 1.2), which has been introduced and refined by Salmon (Section 1.2.1) and further developed, for instance, by the New Mechanists (Section 1.2.2). In Section 1.3 I address the question of what it means when proponents of the CM model call their conception of explanation "ontic". The answer to this question is important to me since I characterize my own account of explanatory reduction as ontic, too, and, in doing so, draw on how this term is understood in the debate about explanation. These introductory notes on CL and CM models of explanation provide the basis on which I can, in Section 2, approach the controversial question of whether the debate about reduction(ism) amounts to a debate about explanations. I argue that this question needs to be answered differently with respect to reduction (Section 2.1) and with respect to reductionism (Section 2.2). Concerning reduction the answer will be negative. In Section 2.1.1 I show why the question of what constitutes the reductive character of biological explanations (the question of reduction) does not boil down to the question of what characterizes an adequate explanation (the question of explanation). In line with this, I argue that my account of explanatory reduction remains uncommitted with respect to whether the CL or the CM model adequately captures the nature of explanation. Concerning reductionism I will give a positive answer, that is, I will argue that disputes about the correctness of explanatory reductionism in fact amount to specific disputes about explanation (Section 2.2.2). Which disputes about explanation these are is specified in Section 2.2.1. One of my major theses will be that one's stance on explanatory reductionism is highly affected by one's stance on the pragmatics of explanation. Thus, in Section 2.2.3 I address the question of whether and in which way pragmatic factors influence the adequacy of explanations. I conclude by addressing the question of whether there can be just one right level of explanation, irrespectively of the research context in which an explanation is developed (Section 2.2.4).

In *Chapter V* I develop my own account of explanatory reduction in biology. I will proceed in two major steps. The *first part* of this chapter (Section 1) serves to specify four concepts that occupy center stage in my analysis of the features of reductive explanations, namely the concept of a biological system, the concept of its context or environment, the concept of parts of a system, and the concept of levels of organization. The specification of these concepts is essential because they are central to answering my main question of which features of biological explanations determine their reductive character. This becomes apparent if one considers the answer I give to this question. My central thesis in the second part of this chapter will be that reductive explanations in biology exhibit three major characteristics: first, they refer exclusively to factors that are located on a lower level than the system whose behavior is to be explained, second, they focus on factors that are internal to (i.e. genuine parts of) the system in question and ignore or simplify environmental factors, and third, they appeal only to the system's parts in isolation. This answer is convincing only if it is clear what a biological system is, how a system is distinguished from its environment, under which conditions an entity counts as a part of a system, and what a level of organization is. I will begin with specifying the concept of a biological system and of its context (Section 1.1). My central thesis will be that biological systems are identified and separated from their environment with respect to the behaviors they typically display (Section 1.1.1) and by identifying the spatial boundary of the system, which ideally is a constant, continuous surface (Section 1.1.2). Finally, I discuss the question of whether the drawing of the system's boundary is a flexible or even an arbitrary matter (Section 1.1.3). In Section 1.2 I clarify what it means to decompose a system into its parts. I distinguish three different notions of parts and argue in favor for one of them (Section 1.2.1). Moreover, I examine the ontological nature of parts (Section 1.2.2) and discuss the controversial issue of whether parts do exist prior to identifying and investigating them (Section 1.2.3). In Section 1.2.4 I introduce different kinds of decomposability (i.e. simple, near, minimal, and non-decomposability) that correspond to different types of biological systems (i.e. aggregative, component, integrated, and non-decomposable systems). Section 1.3 serves to specify the notion of level of organization that underlies my account of explanatory reduction. I critically discuss different notions of levels that have been proposed so far, namely the view that levels of nature correspond to levels of science or theories (Section 1.3.1), Wimsatt's thesis that levels of nature are local maxima of regularity and predictability (Section 1.3.2), and the mechanistic concept of levels (Section 1.3.3). I conclude by delineating the notion of level that I adopt (Section 1.3.4).

On the basis of these conceptual clarifications I can then in the *second part* of this chapter (Section 2 to 6) develop my account of explanatory reduction in biology. I will do so by analyzing, on the one hand, paradigmatic and important examples of reductive explanations from different biological fields and, on the other hand, discussions about reductionism in the contemporary biological literature. The result of this analysis will be that reductive explanations in the biological science exhibit three characteristics: they appeal only to lower-level factors (Section 2), they focus on factors that are internal to the system of interest (Section 3), and they describe only the parts of a system in isolation (Section 4). The first and third characteristics are necessary conditions; the second is only a typical feature that most reductive explanations possess. Taken together, they may be sufficient for an explanation to be reductive. The first feature, that is, the lower-level character of reductive explanations, will be specified in Section 2. I start with analyzing examples from the biological discipline that is regarded as the embodiment of the success of reductionism, namely molecular biology (Section 2.1). In Section 2.2 I point out that reductive explanations are characterized by a unidirectional flow of explanation from the lower to the higher level (Section 2.2.1). Moreover, I spell out what exactly it means for an explanation to refer solely to "lower-level" factors (Section 2.2.2). Section 2.3 introduces two significant subtypes of lower-level explanations, namely fundamental-level or molecular explanations (Section 2.3.1) and single-factor explanations (Section 2.3.2). My central claim in this section will be that although molecular explanations are an important subtype of reductive explanations, it is important to note that reductive explanations in biology are not restricted to molecular explanations. In Section 3 I investigate second characteristic of reductive explanations, namely their internal character. I explicate what it means for an explanation to focus on factors that are internal to the system in question (Section 3.1) and I point out what distinguishes the internal from the lower-level character of reductive explanations (Section 3.2). Finally, I argue that to focus on internal factors implies either completely ignoring environmental factors or simplifying them to a large extent (Section 3.3). In Section 4 I analyze the third feature of reductive explanations, namely that they explain the behavior of a system by appealing exclusively to the system's parts in isolation. The question that I address in Section 4.1 is how exactly the phrase 'parts in isolation' can be understood. My main thesis will be that reductive explanations do only appeal to those relational properties of and interactions between parts that can be discovered by investigating the parts in other contexts than *in situ*. Section 4.2 clarifies the relation between this claim and the assumption that a reductionist treats biological systems as if they were aggregative systems or nearly decomposable systems. I conclude by pointing out the similarities and differences between three kinds of explanation: reductive, part-whole, and mechanistic explanations (Section 5). Furthermore, I explain why I call my account an *ontic* account of explanatory reduction (Section 6).

I Meta-philosophical Preliminaries

> *"If we can understand the science from the inside while retaining a philosophical perspective, we can gain a new and important viewpoint on scientific practice."*
>
> (William C. Wimsatt 2007, 27)

> *"If you wish to learn from the theoretical physicist anything about the methods which he uses, I would give you the following piece of advice: Don't listen to his words, examine his achievements."*
>
> (Albert Einstein 1933, 5)

Content

My aim in this book is to understand an important element of contemporary biological research practice[1], namely reduction. The central questions of this book are, What is reduction in biology? (that is, Which important characteristics of actual cases of reduction can be identified?), and Where do the strengths and limitations of reduction in current biological practice lie?[2] For reasons that I will reveal in detail later on (see Chapter II), my analysis of reduction will focus on examining examples of reductive (and non-reductive) explanations that play an important role in contemporary biology.

By choosing "understanding reduction in actual biological practice" as the goal of my investigation, I have implicitly made some preliminary decisions that will shape the result of my analysis of reduction. In this chapter, I try to make these preliminary assumptions explicit by specifying how I proceed in developing my account of reduction (before developing this account in subsequent chapters,). Questions like this, that is, questions about how to do philosophy of science are located on a *meta-level* of philosophical analysis. Here the question is not what the details of a convincing account of reduction are, but rather what the aim of and what an adequate procedure for generating such an account[3] is. Meta-philosophical questions that are relevant in the context of reduction are for instance: Is the goal pursued by offering an account of reduction to capture actual biological practice or is it to propose an ideal of reduction that need not be realized in practice? What are the corresponding criteria of adequacy for an account of reduction? What role should empirical information about actual biological practice play in the analysis of reduction? What kind of empirical information is crucial – information about how the term 'reduction' is actually used in biology, paradigmatic examples of reductions, biologists' discussions about reductionism, information about what biologists actually do, or information about what biologists themselves think reduction is? Answering these and related meta-philosophical questions is not a mere matter of self-reflection that could be discarded. Rather, this chapter is an *integral part* of my work since it explicates and justifies the philosophical methodology by which I develop my account of explanatory reduction.

Although there is a long-standing debate (labeled 'meta-philosophy')[4] about how philosophy in general is to be carried out, the specific question of how to do philosophy of

---

[1] I use the term 'biological (research) practice' not in opposition to biological theory, but rather in its broadest meaning. According to this, biological practice comprises everything a biologist is typically engaged in when doing research: narrowly "practical" elements (e.g. applying techniques, instruments, or investigative strategies) as well as elements which may be characterized as more "theoretical" (e.g. building theories, modeling, or developing explanations).

[2] What I mean by the "strengths and limits" of reductions and why they are important for my analysis will become more intelligible in the course of this chapter.

[3] In the following I will use 'account' synonymously with 'theory' or 'model' in order to refer to a bunch of philosophical theses about a certain element or feature of a particular scientific field (or science in general).

[4] Questions that are discussed in meta-philosophy are for example: Are there genuine philosophical methods like conceptual analysis, thought experiments, etc.? If yes, what distinguishes these from methods in the natural sciences? Which role does or should a priori knowledge play in philosophy? To what extent can or should empirical knowledge about the natural world be included in the development of, for instance, metaphysical accounts?

science is rarely extensively debated or pursued in its own right.[5] Rather, the meta-philosophical remarks one *can* find are mostly located in introductions of monographs, in which the prior aim of the author is to argue for certain theses in first-order philosophy of science. Thus, only a few of the meta-philosophical discussions in philosophy of science play more than a propaedeutic role. However, this is not to say that there exist *no* interesting and useful theses about the proper aims and methods of philosophy of science in the literature. To the contrary. For instance, William Wimsatt devotes a whole chapter of his book "Re-Engineering Philosophy for Limited Beings" to characterizing the stance and outlook of a "scientifically informed philosophy of science" (2007, 26). Likewise, in the first chapter, "Outlines of a New Philosophy of Science", of the book "Mental Mechanisms" (2008) William Bechtel reveals the naturalistic character of his approach and argues why philosophy of science should be an examination of the "actual practices of science" (2008, 9) rather than a normative endeavor that yields recommendations for future practice. Finally, C. Kenneth Waters devotes an entire paper to the question "What Concept Analysis in Philosophy of Science Should Be" (2004), in which he – contrary to Bechtel – emphasizes the normative character of philosophy of science. One of Waters' main claims is that philosophy should help us to understand how the sciences work (and don't work) with respect to epistemic virtues that we value (2004, 48).

My primary aim in this chapter is not to examine and critically discuss all these different views about how philosophy of science in general should be pursued and, in the end, settle upon one of them. Rather, I review the debate from the perspective of my own approach and try to disclose the meta-philosophical assumptions that lurk behind my analysis of reduction in biology. To do this, I think, is important since it provides clarity to the goals I pursue and to the development of my account of reduction. In what follows I will characterize my account as descriptive (and bottom-up) but critical, as universal but also specific, as including certain normative elements and as being potentially useful for science. But instead of claiming that this is the only game in town, I allow that other kinds of analysis of reduction (involving different aims and different criteria of adequacy) can be adequate as well (see section 2)[6].

In order to give this positive meta-characterization of my account of reduction it is necessary to spell out, for instance, what makes a philosophical analysis of a certain concept (e.g. the concept of explanation, function, causation, or gene) "descriptive", "bottom-up", or "naturalistic", what it means for a philosophical account to be jointly descriptive *and* normative (e.g. Mitchell 2009, 4; Craver 2007a, vii; Woodward 2003, 7), what constitutes the critical character of a philosophical account, in consideration of which empirical information the "description" or "reconstruction" of scientific practice is carried out, and what the alternative to descriptive philosophy of science is. These and other questions are *urgent*

---

[5] This is particularly true with regard to the question of how to develop an account of reduction in biology.
[6] Section references refer to sections in the current chapter.

questions in the meta-philosophy of science that have not been satisfactorily and systematically explored so far. Hence, my second goal in this chapter is to contribute to filling this gap by introducing useful distinctions and clarifying relevant relationships. In so doing I hope to provide a fruitful contribution to the meta-debate about how philosophy of science in general can and should be pursued.

I start my metaphilosophical analysis by pointing out what it means to attempt to understand reduction in current biological research practice (Section 1). One of my main theses is that this aim commits you to focus on cases of reduction that actually occur in biological practice (reduction *in practice*). In other words, it commits you to accept descriptive adequacy as an important criterion of adequacy for your analysis (Section 1.1). I clarify in which sense descriptive accounts in philosophy of science can be called naturalistic (Section 1.2) and why the best way to develop them is to conduct a philosophical analysis in a bottom-up fashion (Section 1.3). I conclude by specifying what the empirical data are that a descriptive account in philosophy of science needs to capture (Section 1.4).

In Section 2, I approach a possible objection which proponents of descriptive accounts (like me) encounter: why care about biological practice in the first place? As a response, I distinguish descriptive projects from normative projects and concede that both aim at different goals and take different criteria of adequacy for granted, which is why both can be accepted as legitimate but different projects. Moreover, I show that there exists a continuum between purely descriptive and purely normative projects and that the projects which are located on the middle ground of this continuum can be distinguished by revealing that they satisfy different dimensions of normativity.

I elaborate one of these dimensions of normativity by arguing that descriptive projects in philosophy of science cannot be *purely* descriptive since philosophers must take up a critical stance on the empirical data scientific practice provides them with (Section 3). In particular, this means that philosophers should, first, develop their account on the basis of examples that are paradigmatic and important, second, that they should explicate assumptions that are only implicitly present in scientific practice, and, third, that they should aspire to construct a coherent account. On the basis of these new insights, I revise the criterion of descriptive adequacy that was presented in Section 1.1.

Section 4 serves to spell out the demand of coherence (i.e. the third respect in which a descriptive account must be critical) by addressing the question of how much and which kind of pluralism we in fact need. After telling apart two different senses of non-universality (Section 4.1), I propose a second criterion of adequacy that can be condensed into the motto "Try to achieve as much generality as you can get and as much specificity as you need" (Section 4.2). Finally, I discuss the objection that this criterion is not pluralistic enough (Section 4.3).

In Section 5 I pick up on the issue of normativity again. The central task of this section is to clarify what it means when some philosophers of science state that their account is jointly descriptive and normative. Several authors link the normativity of their project to the fact that it aims at disclosing certain epistemic norms, such as the standards according to which mechanistic explanations in neuroscience are assessed as adequate or inadequate or as good or bad (Craver 2007a). In order to clarify the notion of normativity that is involved here I first specify the concept of epistemic norms in science (Section 5.1). Then I point to an important distinction, namely the one between describing epistemic norms that are in fact accepted in science and positing epistemic norms that should apply to science (Section 5.2). I argue that several philosophical accounts are normative in the sense that they are critical descriptions of existing norms (which is the second dimension of normativity), not in the sense that they provide independent justifications of epistemic norms (which would be the third dimension of normativity). In Section 5.3 I claim that those philosophers who seek to justify which norms should apply to science face a dilemma: either they commit an is-ought fallacy or they fail to specify the grounds on which their independent justifications are made. I conclude by pointing out that my own analysis of reductive explanations in biology is normative only according to the second dimension of normativity (Section 5.4).

In the last section of this chapter (Section 6) I address the popular question whether philosophy of science should be judged according to its relevance to science (and to the broader community). I argue that it is implausible to regard actual utility as a measure of the quality of a philosophical account. However, this does not preclude that a philosophical account should be potentially useful to scientific practice. This constitutes the fourth and last criterion of adequacy for an account of reduction that aims at capturing what reduction in contemporary biological practice actually is.

1 Describing Biological Practice

1.1 Descriptive Adequacy

As I stated at the beginning of this chapter, my aim in this book is to understand what reduction in current biological research practice is and where the strengths and limits of performing reductions lie. To pursue this goal presupposes *takeing biology seriously*, that is, focusing on the analysis of cases of reduction that are actually present in and that are crucial to biological practice. Philosophers who aim to understand biology are interested in cases of reduction that are realized in biological practice (reduction *in practice*), not in an ideal of reduction that can only be achieved in principle (reduction *in principle*). Thus, by identifying "understanding biological practice" as the aim of a philosophical account of reduction one accepts the following criterion of adequacy (which is further refined in Section 3):

An account of reduction is adequate only if

(1) it captures cases of reduction that occur in current biological research practice.[7]

---

[7] This holds for other accounts in the philosophy of science, too.

The kind of adequacy that is captured by this criterion is also called *descriptive adequacy* (e.g. Craver 2007a, 19; Machamer, et al. 2000, 8).[8] The underlying idea is that the analysis of a certain element of the sciences (e.g. explanation, prediction, the concept of a gene, or reduction) can only help to understand actual biological practice if it captures for instance those explanatory strategies that are typically pursued in biology, if it accounts for real cases of successful or failed predictions, if it copes with how the term 'gene' is *de facto* used in different biological fields, and if it captures those strengths and limitations of reductive strategies contemporary biologists highlight. In other words, an adequate account in the philosophy of science must "*save the phenomena*" (about scientific practice), to borrow a phrase from Bas van Fraassen (1980, 41). Granted, this criterion is quite general and one might accuse it of being too unspecific. That is why, in Section 3, I further specify what it means for an account to be descriptively adequate.

1.2 Methodological Naturalism

Many philosophers of science rarely speak about the descriptive character of their account. Rather, they disclose their affiliation to the naturalistic tradition in philosophy (e.g. Bechtel 2008, 4-10). This raises the question of why an account in the philosophy of science is called naturalistic on the grounds of its descriptive dimension.

The concept of naturalism (just as the concept of reduction) has a long history, is multifaceted, and is not easy to specify. However, in the debate it is quite common to identify *methodological naturalism* as an important kind of naturalism. It seems to me that this is exactly the kind of naturalism that comes into play in the case of pursuing philosophy of science in a descriptive manner. According to David Papineau a methodological naturalist "see[s] philosophy and science as engaged in essentially the same enterprise, pursuing similar ends and using similar methods" (2009). Although the details of how methodological naturalism is characterized vary (e.g. Plantinga 1996; Keil/Schnädelbach 2000; Kornblith 2007), the general statement about philosophy remains largely the same: with respect to their methods (and aims) there exists no principled difference between philosophy and the natural sciences. That is, philosophy can be pursued by applying methods that are similar to those successfully employed in the natural sciences (and vice versa).

Although – with respect to philosophy of science – I feel the attraction of this thesis, I agree that it leaves much room for criticism. For instance, one could allude to the vagueness of the term 'similar' and point out that in some respect any two methods can be similar to each other. Or one could object to methodological naturalism by highlighting the methodological differences between philosophy and the natural sciences. One could, for example, emphasize the importance of *a priori* considerations and intuitions for developing

---

[8] A related criterion of adequacy is what Love calls "epistemic transparency", which demands "a descriptive correspondence between philosophical theories about science and scientific practice" (2012, 179).

philosophical theses and oppose this to the natural sciences where empirical investigations like experiments play a crucial role. A methodological naturalist could, in turn, counter that there is no such clear distinction between the methods of philosophy and those of the natural sciences since philosophy also makes use of empirical data and experiments (e.g. thought experiments) and the natural sciences are not free of a priori considerations. In this chapter I do not want to decide which of the two is right. As it seems to me the thesis that philosophy of science can – independently of the goal one pursues – only adequately be carried out in a naturalistic fashion (i.e. by applying methods that are similar to those in the natural sciences) is unnecessarily strong. Thus, I leave it open whether philosophy of science really is "part and parcel of… science itself" (Rosenberg 1985, 2) or whether it can be methodologically distinguished from science.[9]

What I want to call attention to in this section is that there exists in fact a similarity between conducting philosophy of science *in a descriptive fashion* and performing research in the natural sciences. And exactly *this* is the reason why many philosophers of science characterize their descriptive account as naturalistic. The similarity between descriptive philosophy of science and the natural sciences is that in both cases we have "empirical data" or "phenomena" that need to be captured (or "saved") by the account or theory that is developed.[10] The difference is that in the natural sciences these empirical data are data about the natural world itself,[11] whereas in philosophy they are data about the natural sciences (e.g. about explanations scientists develop for natural phenomena, about methods by which scientists investigate these phenomena, or about the causal inferences scientists make).[12] In other words, philosophy of science is a *second-order discipline* (Carrier 2007, 15) and the phenomena it tries to account for are also located on a "second level". That is, the empirical data against which a philosophical theory is tested are data about the natural sciences, which in turn develop scientific theories that are tested against data about the natural world.

However, the question remains how far this similarity between the methodology of descriptive philosophy of science and the natural sciences goes. In order to resolve this, we need to further elucidate how descriptive and naturalistic philosophy of science is carried out. In this context two questions need to be answered: one concerns the process of description, the other the phenomena to be described. First, how does theory building (e.g. the construction of an account of reduction) in philosophy of science proceed? Second, what are these empirical data about the sciences that an adequate (naturalistic) account needs to capture – the actual usage of a term, clear-cut and uncontroversial examples, self-reports of

---

[9] Even if one argues for a methodological continuity between philosophy and the natural sciences one can still claim that a philosophical enterprise can be clearly distinguished from a scientific enterprise, for instance by pointing out that they pursue distinct aims and seek to answer different questions.

[10] Thus, we could also speak about *empirical adequacy* instead of descriptive adequacy.

[11] Presupposing that some kind of scientific realism is true.

[12] Callebaut adopts a similar view: "The naturalistic perspective implies that *matters of fact* are as relevant to philosophical theory as they are relevant in science." (1993, 1)

scientists, or philosopher's observations of "the scientific practice"? Let us start with investigating the first question, that is, with exploring the process of description by which a philosophical account is generated.

1.3 Bottom-up Philosophy of Science

As I pointed out before (see Section 1.1), to pursue the aim of understanding contemporary science commits one to taking actual scientific practice seriously and developing an account that is descriptively adequate. I think this requires that another, closely related condition is satisfied: Philosophy of science should be "*bottom-up*" (Bickle 2003, 31; my emphasis), that is, it should emerge from a detailed investigation of contemporary scientific practice. Philosophers should "watch… science at work" (Macilwain 2009, 840) and even not "be afraid to *do* science" (Wimsatt 2007, 26) if this is necessary for gaining important insights into how science is actually carried out. That a philosophical account is developed in a bottom-up fashion means that one starts with a detailed examination of scientific practice. For example, if one wants to generate an account of reduction, one should begin for instance with the study of a wide range of examples for reductive explanations, with the analysis of discussions about reductionism that can be found in scientific research papers, or with an investigation of reductive strategies prevalent in scientific practice.[13] The goal is to understand the practices of science in an "internal" (Bickle 2003, 32) manner or "from the inside" (Wimsatt 2007, 27) while, at the same time, retaining the "interpretive distance" (Love 2008a, 67) that is necessary for a philosophical perspective.

The opposite procedure of developing a philosophical thesis about science can be called "*top-down*". It includes for instance the search for confirmation of a general philosophical thesis in a certain scientific field. When philosophy of biology emerged as a separate discipline in the 1960s and 70s this way of conducting philosophy of biology was wide-spread. Most notably, the effort to apply Nagel's model of theory reduction to biology (more precisely, to the relationship between Mendelian genetics and molecular biology) marks the origin of the reductionism debate in the philosophy of biology (cf. Schaffner 1967, 1969; Hull 1974, Kitcher 1984). As Paul E. Griffiths puts it in a nutshell: "biological science [is used] as a *testing ground* for claims in general philosophy of science" (2007, 69; my emphasis). Such a top-down approach is problematic in so far as it can involve the danger that one imposes an ill-fitting, normative model on science. In addition, a top-down procedure can tempt one to distort the empirical basis for example by selecting not the phenomena that are paradigmatic and important but that support the philosophical theory. By contrast, developing a philosophical account in a bottom-up fashion ensures that it captures what is characteristic of and what is crucial to real science.[14]

---

[13] As Carl F. Craver once recommended to me: "You should start by letting science speak to you."

[14] However, I do not want to assert that it is in principle impossible to develop a descriptively adequate account in a top-down manner. My claim is that (presupposing the aim of understanding actual science) philosophy of

1.4 The Data to Be Captured

Let us now turn toward the second of the two questions presented above (see Section 1.2): What are the empirical data that an adequate (naturalistic) account in the philosophy of science needs to capture? What does "watching science at work" (Macilwain 2009, 840) exactly mean?

A first clue can be found in the tasks that are typically assigned to philosophy of science (or philosophy of biology in particular). Most authors agree that a significant job for philosophers of science is to analyze concepts that are central to science (e.g. explanation, progress, model, complexity, law, experiment, mechanism, etc.), including concepts that are used as technical terms in the sciences itself (e.g., with respect to biology, fitness, ecological niche, gene, innateness, biodiversity, etc.). By providing conceptual clarification philosophers can also contribute to the clarification of the structure of existing problems or questions as well as to the specification of what is required to solve a problem and what counts as a satisfactory answer to a question (Love 2008a, 72; Rosenberg/McShea 2008, 4).

However, this still leaves open the question of which empirical facts are to form the basis upon which the meaning of scientific concepts it clarified. One possibility is to study how a term is actually used by scientists, as is done by Karola Stotz and Paul E. Griffiths (2004, 2005) in their extensive studies of the actual usage of the gene concept in biology (called "Representing Gene Project (RGP)"). But are empirical studies that record self-reports of scientists about what they think a certain concept means really the appropriate empirical basis of a conceptual analysis? Some philosophers have challenged this. They argue that the self-perception of scientists offers only limited insight into how a concept is actually used since the picture scientists have of their own work and their actual behavior often come apart (e.g. Carrier 2007, 15f; Falkenburg 2005, 92).[15] However, to draw upon the actual usage of a concept while analyzing its meaning does not commit one to rely on the results of surveys under scientists. The different kinds of reasoning practices (e.g. the explanations of phenomena in a certain field, the inferences that are typically drawn, the theoretical assumptions of certain modeling techniques, controversial disputes about certain research topics, etc.) seem to be a much better source for detecting the actual usage of a concept.

As some philosophers have convincingly pointed out, how scientists use words cannot be the only empirical basis on which a philosophical account is developed and against which its adequacy is tested. A conceptual analysis should not only capture what scientists *say* but also what they *do*, that is, how they work. James Woodward, for instance, emphasizes that his interventionist theory of causation goes beyond being a "mere"

---

science should be pursued in a bottom-up fashion since this guarantees that the resulting account is sensitive to real science.

[15] This is not to say that biologists' reflections about their own discipline - like Ernst Mayr's "What Makes Biology Unique?" (2004) or Francis Crick's "What Mad Pursuit: A Personal View of Scientific Discovery" (1988) – are not valuable for philosophers.

conceptual analysis that describes the actual usage of the term ‘cause’ (2003, 7). One reason he cites is that he focuses “not just on how people use words, but on larger practices of causal inference and explanation…, practices that involve substantial non-verbal components” (2003, 7).

Likewise, I seek to capture a wide range of empirical phenomena in my analysis of the concept of reduction (or, more precisely, of the concept of reductive explanation in biology). That is, in developing my account of reduction I consider typical and clear-cut examples of reductive explanations from different biological fields, I analyze how biologists evaluate the strengths and limits of reductive strategies (or, as they say, the “limits of reductionism” Ahn, et al. 2006a, 709 and Mazzocchi 2008, 10) in biological research practice, I explore how these reductive strategies are actually carried out, and I examine how the adequacy of reductive explanations is assessed in the life sciences.[16]

2 A Continuum between the Descriptive and the Normative

2.1 Descriptive and Normative Projects in Philosophy of Science

An opponent of my descriptive, naturalistic, and bottom-up way of analyzing reduction in biology could object: why care about biological research practice in the first place? Why not regard a philosophical account of reduction as an *ideal* of how science *should* proceed or, in the long run, *will* proceed (when seeking reductions)? He could insist that this philosophical ideal need not be realized in part or in whole in contemporary biological practice. And he could point to some important predecessors who have already adopted this strategy. For instance, Kenneth F. Schaffner admits that his general reduction-replacement (GRR) model is only “*peripheral*” (1974a, 111; 1993, 509; my emphasis) to biological practice since molecular biologists are not interested in obtaining the “complete chemical characterizations” (1974a, 127) that are, according to the GGR model, required for theory reductions.

> [S]uch [physicochemical] detail is not of primary interest to molecular biologists, and they will rarely orient their research in such a way as to attempt to provide such extensive detail. (Schaffner 1974a, 128)

In spite of the peripherality of his GRR model, Schaffner clings to it as a regulative ideal that should guide the development of molecular biology (1993, 511).[17]

---

[16] It should have become clear that what I and other philosophers of science (e.g. Waters 2008) mean by ‘conceptual analysis’ deviates from the notion of conceptual analysis characterized by Frank Jackson (1998; Chalmers/Jackson 2001). The method of conceptual analysis as an *a priori* analysis of our philosophically interesting everyday concepts and folk theories is also known as “the Canberra plan”.

[17] In his recent work Schaffner has taken up a more compliant stance. For instance, in his paper on “Reduction: the Cheshire Cat Problem and a Return to the Roots” (2006), Schaffner concedes that “what have traditionally been seen as robust reductions of one theory or one branch of science by another more fundamental one are

It seems to me that there are two options that need to be distinguished: on the one hand, one can develop an account of reduction that aims at capturing and understanding actual biological practice and that proceeds by describing this practice (descriptive project). On the other hand, one can (as Schaffner once did and David Lewis (1994) always did) give an account of how the concept of reduction should be understood without caring about what cases of reductions actually performed in biological practice look like (normative project). I think these two options are best seen as completely *different projects*. Those philosophers who want to understand what biologists actually do and how biological research practice really works will not be satisfied with a philosophical account that merely reflects the ideals of philosophers but does not capture what is really going on in biology itself. They will judge accounts of the second kind as descriptively inadequate and, probably, not continue thinking about them at all. Philosophers who pursue a project of the second type (i.e. a normative project) do not share the goal of understanding actual biological research practice, but rather endorse other aims and values of a philosophical account. For example, they seek to develop an account of reduction that captures certain philosophical or common sense intuitions, that fits well into a broader philosophical theory (e.g. into a certain metaphysical picture of the world), that is universally applicable, or that has special explanatory force.[18] In the extreme version of this kind of project, descriptive adequacy is simply abandoned as a criterion of adequacy. The focus lies exclusively on analyzing reduction *in principle*, that is, on saying what reduction "really" is. What characterizes reduction *in practice* is ignored.[19]

descriptive project	**normative project**
aim: understanding biological practice (reduction *in practice*)	aim: explicating reduction *in principle*
criterion of adequacy: descriptive adequacy	criteria of adequacy: accounting for intuitions, metaphys. suitability, etc.

Figure I.1: Descriptive and Normative Projects in Philosophy of Science

---

largely a myth" (2006, 378). In the biological sciences we typically find "creeping reductions" (i.e. partial, multi-level reductive explanations) instead of "sweeping reductions" (2006, 397). However, although this seems as an immense departure from his original position, many details of Schaffner's recent work convey that he is still influenced to a great extend by his GRR model. For further details see Chapter 2.

[18] Love identifies further possible criteria of adequacy: "universality" (2012, 175), "past relative success", "having a source of normativity" (2012, 180), "aiming for more abstract… formulations of reasoning", and "providing an account of (scientific) rationality" (2012, 182).

[19] This difference between analyzing reduction *in practice* and developing an account of reduction *in principle* does *not* fully coincide with the difference between in practice versus in principle claims in the reductionism debate (see Chapter II, Section 1). Even if you pursue a descriptive project and focus on the analysis of cases of reductions that are actually carried out *in practice*, you can, nevertheless, utilize the understanding of reduction you get to argue for *in principle* reductionism (e.g. for the thesis that, in principle, all biological explanations can be reduced to molecular explanations; Rosenberg 2006). However, if you think it is important to reconstruct actual cases of reductions in order to understand what reduction is (reduction in practice) it is likely that you will restrict your reductionist or antireductionist claims to cases of reductions that can actually be accomplished at present (in practice reductionism).

What is crucial to note is that I do not claim that projects of the second kind (i.e. normative accounts of reduction) are misleading or do not instantiate "healthy philosophy" (Kitcher 2011, 249). In other words, I do not go along with the methodological naturalist who states that *any* project in philosophy of science must be conducted as a descriptive project (see Section 1.2). I concede that there are more games in town (or more ways to play the game in town) than the descriptive-naturalistic enterprise. What I contend is that *if* one shares the aim of understanding contemporary scientific practice, *then* one needs to accept descriptive adequacy as an important criterion of adequacy. More specifically, I claim that if one wants to understand what reduction in biology actually is and where the advantages and limitations of performing reductions in current biological practice lie, then one should develop an account of reduction in a descriptive manner.

2.2 Refining the Picture: Dimensions of Normativity

However, the way philosophy of science is presently carried out reveals that the situation is more complicated than the picture just presented suggests. Let me elaborate on two points. First, it must be noted that in contemporary philosophy of science *purely normative* projects, in which facts about how science is actually undertaken are treated as completely irrelevant, are rare. That is, few accounts in current philosophy of science include *no* empirical information about actual scientific practice at all. This has not always been the case. For instance, in the heyday of Logical Empiricism many authors endorsed formal theories of science, which frequently were purely normative projects.[20] But now, especially in philosophy of biology, the majority of projects belong to the descriptive kind (even if not to the purely descriptive kind). Philosophers want to understand, for example, how the success and failure of explanation in biology is in fact evaluated, why molecular research is as important as it is, which different roles models play in biological research practice, and how biologists *de facto* estimate the scope of biological generalizations. However, as I said before, there exist philosophers of science who are not primarily interested in capturing and understanding actual scientific practice. Their goal is to develop a view about science or about a specific element of science (like explanation, causation, confirmation, law, etc.) that is adequate, for instance, because it captures certain philosophical or common sense intuitions, because it is in line with a certain metaphysical picture of the world, or because it has special explanatory force. But even the projects of this kind are rarely pursued in a *purely* normative manner, that is, without including (at least partially) certain empirical claims about how science really works and why it is actually successful.[21] For instance,

---

[20] For example, with regard to his account of scientific explanation Hempel emphasizes that it is "not meant to describe how working scientists actually formulate their explanatory accounts" (1965, 412).

[21] This is not surprising since it seems weird to make claims about how science ideally should work or how certain elements of scientific practice like explanation and reduction should be understood without taking into account how science actually works and what scientific explanations and reductions *in fact* are.

although Schaffner's GRR model of reduction clearly is a normative account (since it is not developed by reconstructing actual cases of reductions, but posits how reductions ideally look like) it is, nevertheless, illustrated by examples from biology (see Schaffner 1993, 432-487). The same applies to Brandon's account of adaptation explanation in evolutionary biology (1990, 159-194). In line with Schaffner, he argues:

> Although very few, perhaps one or two, adaptation explanations in evolutionary biology meet this standard, it is, I have argued, a useful *normative ideal*. For any proffered adaptation explanation it provides a checklist that will show where more explanatory information is needed. (Brandon 1996, 197; my emphasis)

Despite the obvious normative character of his account, Brandon does not completely ignore those adaptation explanations that are in fact proposed by contemporary evolutionary biologists. However, empirical information about real adaptation explanations does not play a noteworthy role in the development or justification of Brandon's account. It seems as if Schaffner as well as Brandon employs the examples for mere illustrative or refining purposes. Thus, we can conclude that what makes a philosophical project normative is that it does not aim at understanding actual scientific practice (i.e. it does not accept descriptive adequacy as a criterion of adequacy). And if it includes empirical information about scientific practice then this information plays no central role in developing or justifying the results of this project.

Second, descriptive projects in philosophy of science are far from being carried out in a *purely descriptive* manner. Rather, descriptive projects are frequently said to give rise to or to contain normative claims about how science should be carried out or how a certain scientific concept should be understood (see, for instance, Mitchell 2009, 4; Craver 2007a, vii; Waters 2004; Woodward 2003, 7). One of my main theses in this chapter is that, in addition to the already introduced notion of normativity that distinguishes normative from descriptive projects, two further dimensions of normativity need to be kept apart. On the one hand, descriptive projects in philosophy of science inevitably involve critical elements that allow for distinguishing philosophical approaches from mere descriptive sociological, historical, or psychological ones.[22] These critical elements give rise to a philosophical methodology that can be characterized as critical-normative (more on this in Section 3). On the other hand, a philosophical project can be both descriptive and normative at the same time. That is, it aims at understanding scientific practice; in particular, it aims at reconstructing the epistemic norms that are *in fact* accepted in science. But at the same time or after having reached this goal, projects of this kind involve an independent justification of those epistemic norms that *should* be accepted in scientific practice. This third dimension of normativity will be examined in Section 5.

---

[22] To be clear, I do not claim that *all* sociological, historical, or psychological projects are purely descriptive, just that some of them are and that no philosophical projects are purely descriptive.

These three *dimensions of normativity* can be summarized as follows. Philosophical accounts can be characterized as normative because:

(1) their *goal* and *methodology* are normative. That is, they aim at developing an account of how science should ideally be carried out or how a scientific concept should be understood. Accordingly, descriptive adequacy is rejected as criterion of adequacy and empirical information plays no central role in their account.
→ They are called *normative* projects.

(2) their *methodology* is descriptive-critical (i.e. partly normative). That is, they aim at understanding scientific practice and accept descriptive adequacy as criterion of adequacy. However, they hold that the adequate philosophical methodology is not pure description, but rather critical reconstruction (see Section 3).
→ They are called *descriptive-critical* projects.

(3) their *goal* and *methodology* are descriptive and normative together. That is, they aim at understanding scientific practice (in particular, the in fact accepted epistemic norms). But on the basis of this understanding they deliver an independent justification of those epistemic norms that should be accepted (see Section 5).
→ They are called *descriptive-normative* projects.

2.3 A Continuum of Projects

Let me put the things together. On the one hand, there is a crucial difference between what I call descriptive and normative approaches in philosophy of science (first dimension of normativity). Those who pursue a project of the former kind aim at understanding actual scientific practice and accept descriptive adequacy as an important criterion of adequacy. By contrast, those who carry out a normative project abandon this goal and replace descriptive adequacy with other criteria of adequacy (e.g. capturing intuitions, fitting to general philosophical theories, explanatory force). They do not show how scientific practice in fact works, but what it ideally *should* look like. On the other hand, this difference becomes blurred a bit since, in fact, there exists a *continuum* of different projects ranging from purely descriptive projects to purely normative projects. In other words, these two extremes define a whole spectrum of possible projects and sometimes it is quite difficult to decide where the one kind of project ends and the other begins. Descriptive projects involve critical-normative elements (second dimension of normativity; see Section 3). And sometimes philosophical projects aim at understanding scientific practice, but nevertheless contain an independent justification, for example, of certain epistemic norms that should be accepted in science. Accordingly, projects of this kind cannot be neatly assigned to one of the two layers, but rather are characterized as being descriptive and normative at once (third dimension of normativity; see Section 5). Furthermore, some normative projects incorporate empirical information about how science in fact works and is successful. In fact, the majority of accounts in philosophy of science are located somewhere on the *middle ground* of the continuum between purely descriptive and purely normative projects. As I will argue in Sections 3 and 5, my project of developing an account of explanatory reduction in biology belongs to the descriptive-critical kind of projects, not to the descriptive-normative kind. The variety of different projects is illustrated by the following figure:

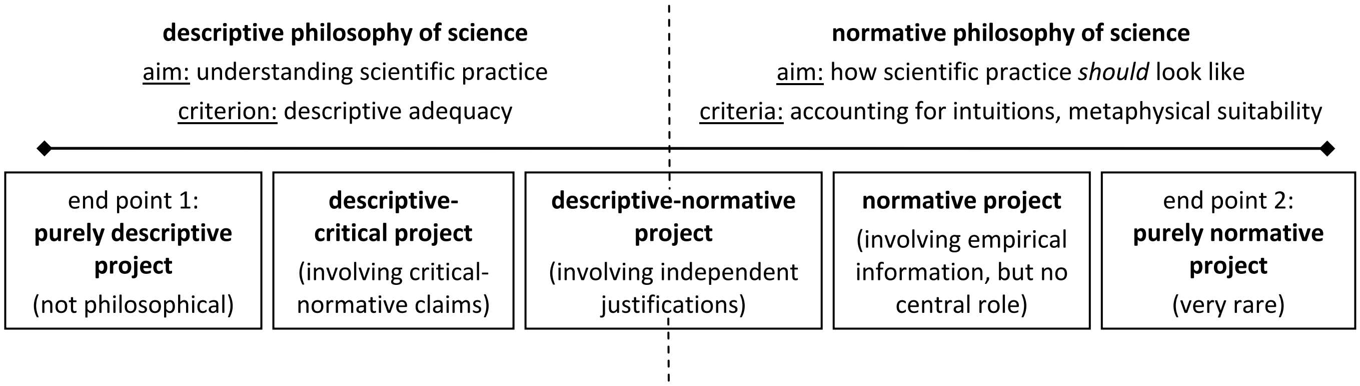

Figure I.2: A Continuum of Different Kinds of Projects in Philosophy of Science

3 Why Pure Description Is Not Enough

Let us now come back to the first of the two questions identified at the end of Section 1.2: How does one proceed in developing a philosophical account about a certain element of scientific practice? What does ‘describing how science in fact works’ exactly mean? And how can the criterion of descriptive adequacy, which was presented in Section 1.1, be further specified? As I have shown in Section 1.3, descriptive accounts in philosophy of science should emerge from a detailed investigation of current scientific practice. But how do we get, for instance, from the particular examples of reductive explanations and from biologists’ discussions about the limits of reductionism to a general account of reductive explanations in biology – do we do it by describing all available cases and perhaps generalizing them?

In this section, I argue that the process of developing a philosophical account of science is to be characterized as an active, *critical reconstruction*[23] or “explication” (Carnap 1950) rather than as a passive description of how science is done.[24] In order to provide clarity and understanding philosophers of science cannot merely passively picture scientific practice one to one. It is simply not possible to read off a certain philosophical account (e.g. what a reductive explanation is, what makes science successful, what makes a trait of an organism to a function, what causation is, etc.) from scientific practice. Thus, philosophers of science must take up a critical stance on the empirical data science provides them with. That

---

[23] The way I understand the term ‘reconstruction’ is similar to Hans Reichenbach’s notion of a “rational reconstruction” (1938, 6) except that I deny that what belongs to a critical reconstruction is determined by the standards of logic alone (see also Waters 2004, 34-38).

[24] This is not a revolutionary view. I suppose that, in fact, most philosophers of science agree with me (see for example Schurz 2005, Carrier 2007, and Waters 2004).

is, they must sort out phenomena that are not typical and not important, try to articulate assumptions that are only implicit in scientific practice, and construct a coherent philosophical account. Since this critical stance involves normative assumptions, for instance, about which examples should be conceived as paradigmatic, how certain concepts should be understood, and which statements of biologists should be included in the empirical basis of an account, this procedure can be characterized as normative (according to the second dimension of normativity, see Section 2). At least three reasons can be given why philosophy of science must go beyond the purely descriptive and exhibit a *critical-normative character*:

First, the analysis of a certain scientific concept cannot and should not capture *all* available data (e.g. all reductive explanations that have been ever developed in each biological field). Rather, it should be focused on the reconstruction of "*paradigmatic*" (Wimsatt 2007, 27) and *important* examples. Since the resources of a philosopher are limited it is simply impossible to analyze all available cases from each scientific field. In addition, such a procedure would be disadvantageous since the heterogeneity of cases would be overwhelming (for detailed argumentation see Section 4). Thus, a philosopher should single out those cases (e.g. of reductive explanations, causal inferences, functional ascriptions, etc.) as the empirical basis of his account that are, on the one hand, typical or representative (in short, paradigmatic) for certain scientific fields and that are, on the other hand, important to these fields. The importance of a certain case to a field can have different reasons. The most general reason is that it is a significant example for how the *success* (or failure) of research in a certain scientific field is promoted. Reconstructing instances of, for example, successful or failed reductive explanations are of particular importance for building an account of reductive explanation since philosophers not only want to understand "how science is done" but also "why it is as successful as it is" (Giere 1999, 53). In this sense the criterion of adequacy that John Norton identifies, namely "successful functioning" (2003, 648), can be characterized as a subtype of the criterion of descriptive adequacy. Other reasons why particular cases are important to a field are that they contribute to achieve a certain aim of the field (for example, some explanations in neuroscience promote the manipulation of the brain; Craver 2007a, ix) or that they are subject to intensive debate in the field. The decisions philosophers have to make about which empirical phenomena are paradigmatic and important cannot be read off scientific practice and, thus, constitute a departure from the purely descriptive endeavor.[25]

Second, the task of philosophy of science is also to *make explicit* assumptions that are only implicitly present in scientific practice (e.g. Mitchell 2009, 4; Carrier 2007, 17; Craver 2007a, x; Love 2008a, 68). This is the second reason why the construction of an account or

---

[25] Some philosophers argue that the indispensable step of sorting out paradigmatic and important examples displays the *normative* character that a putatively descriptive kind of philosophy of science also has (Gesang 2005, 18; Andersson 2005, 76 f; Janich 2005, 155f). I prefer to call this the critical character of philosophy of science in order to clearly distinguish it from the other respects in which philosophy of science can be normative (see Section 1.2).

theory about a certain element or feature of science must involve more than a mere description of scientific practice, namely a critical element. My project of developing an account of reduction (more precisely, an account of reductive explanation) presents an instructive illustration of this need of explication.

If one wants to figure out what it is that makes a biological explanation reductive the first challenge one encounters is that only very few biologists indicate whether the explanations they develop exhibit a reductive character or not. Biologists propose explanations for certain phenomena and they argue about whether the explanations are adequate or not (i.e. whether they succeed or fail). But beyond that they usually do not indicate whether an adequate explanation is reductive or not. This difference is just not important to them. The only exception seems to be the case in which the reductive character of an explanation is the reason for its inadequacy or failure. In these cases biologists engage in intensive debates about the “limits of reductionism” (e.g. Ahn, et al. 2006a, 709; Mazzocchi 2008, 10) and the need to move “beyond reductionism” (Gallagher/ Appenzeller 1999, 79). However, in these debates, too, biologists rarely speak about reductive explanations themselves. Rather, they discuss the correctness of reductionism and the adequacy of the “reductionist stance” (Soto/Sonnenschein 2009, 3) or of applying reductive methods in investigating biological phenomena. Sometimes biologists do not even use terms beginning with ‘reduction…’ but, nevertheless, talk about reduction. This is the case when they speak about explaining the behavior of a system by reference to its parts or about the method of analysis or decomposition. Thus, what biologists regard as a reductive explanation and which conditions determine the reductivity of an explanation often is only *implicit* in biological practice and needs to be unfolded by philosophers.

Against this background, the above presented criterion of adequacy for an account of reduction (see Section 1.1) needs to be refined:

> (Presupposing that the aim is to understand biological practice,) an account of reduction is adequate only if
>
> (1*) (a) it captures the *paradigmatic* and *important* cases of reduction that occur in current biological research practice and
>
> (b) *explicates* the understanding of reduction that is (often only implicitly) present in biological practice.

Third, what biologists mean by ‘reductionism’ and what they take to be the constraints of a reductive explanation is *by no means homogenous* and sometimes it is even *conflicting*. The same holds for other aspects of science. Following Reichenbach (1938), Waters characterizes the actual thinking of scientists as “vague” and “fluctuating” (2004, 38). He agrees with Reichenbach that this calls for a critical element in the otherwise descriptive approach of philosophy of science (Waters 2004, 38-41). Similarly, Woodward recognizes that “causal and explanatory claims [in science] sometimes are confused, unclear and ambiguous" (2003, 7). Thus, in order to provide a coherent and unified account, philosophers need to adopt a critical-normative stance, that is, to identify some cases as clear-cut, some claims of scientists as correct, and others as too vague or as incorrect.

To illustrate this point, consider again my project of developing an account of reductive explanation. As I pointed out before, a large part of the empirical data that are available are research or review papers in which biologists discuss the virtues and limits of reductionism. The challenge one encounters is not only that in these papers biologists rarely explicitly speak about the characteristics of reductive explanations (second point), but also that the claims they make are not homogeneous and sometimes even incompatible with another. Especially when it comes to the exact conditions under which an explanation is judged to be reductive or non-reductive, there is often conformity, but no total agreement. For example, some biologists identify reductive explanation with additive explanations, that is, with explanations in which a biological system is treated as an aggregative system (e.g. Strange 2005, 968; Bizzarri, et al. 2008, 181; Kitano 2002, 1662). Other biologists explicitly reject this claim because it results in a too restricted view of reductive explanation (see Chapter V, Section 4.2). They state, for instance, that "[m]olecular biologists… do not hold the naive view that complex structures and processes are just sums of their parts" (Fincham 2000, 343; see also Wilson 1988, 270). If one wants to develop a coherent account of reductive explanation one needs to ponder which of these claims should be integrated into the account (e.g. because it is a common claim or because it suits best with the other empirical data) and which should be sorted out as being incorrect, rare, too vague, or insufficiently justified. For instance, either the reference to complex forms of organization violates the reductive character of an explanation or not. Even though the usage of the term 'reductionism' of one biologist suggests that the one assumption is correct and the statements of another biologist supports the opposite assumption, in a coherent philosophical account you cannot have it both ways. The need for coherence considerations of this kind constitutes the third respect in which philosophy of science must go beyond passively describing scientific practice and must exhibit what I call a critical character.

To sum up, philosophers of science must take up a critical stance on the empirical data that scientific practice provides them with. Put another way, developing a philosophical theory about a certain element of science requires more than mere description. Rather, the adequate method is critical reconstruction. This means, in particular, that philosophers should, first, develop their account on the basis of examples that are both paradigmatic and important, second, that they should explicate assumptions that are only implicitly present in scientific practice, and, third, that they should aspire to construct a coherent account. This critical stance can also be characterized as normative since it involves normative assumptions[26] (according to the second dimension of normativity, see Section 2).

---

[26] For instance, it involves normative claims about which examples should be conceived as paradigmatic, how certain concepts should be understood, and which statements of biologists should be included in the account.

4 How Much Pluralism Do We Need?

In the previous section I claimed that descriptive philosophy of science must involve a critical element since it is aimed at construing coherent and universal models of science (third point). With respect to this call for universality a pluralistically inclined philosopher could object: Why seek universality at all? He could continue by pointing at the many differences that exist among and within the sciences and argue that philosophy of science must be specific and highly context sensitive in order to be in touch with real science. In this spirit, Bechtel argues that "the naturalist is led to be a pluralist" (2008, 10). By this he means that "the naturalist is committed to developing accounts that work for specific sciences" (2008, 9f). Similarly, William Wimsatt claims that any adequate "meta-theory" about science will have exceptions and be context-dependent, in other words, will be "full of *ceteris paribus* qualifiers" (2007, 30). Both of them seem to have a point. However, it is crucial to note that there are different notions of non-universality (or specificity) involved here. In order to assess how much and which kind of pluralism we need, the following different respects in which a philosophical account about science (or parts of the account) can be non-universal need to be kept apart.

4.1 Two Senses of Non-Universality

First, an account can be non-universal in the sense that it does not hold for science in general. Rather, its *scope of application* is restricted to a specific field of science. For example, one might hold (as I do) that the question what a reduction or reductive explanation is needs to be answered differently with respect to, for instance, biology, physics, or the social sciences (or even with respect to different biological sciences). This is the kind of specificity or pluralism Bechtel has in mind when he states that the naturalist is led to be a pluralist. It seems to me that such a restriction in scope is often very reasonable. Since there exist in fact many differences between the various scientific fields a limited focus often allows a more fine-grained and, thus, descriptively adequate analysis. What the appropriate scope of application is depends, of course, on the philosophical question that is asked as well as on the state of the proposed area of application. Thus, it needs to be determined on a case-by-case basis.

Second, there is another sense in which an account in philosophy of science can be non-universal. Within its range of application (which is restricted or not) the different theses that an account contains can hold either without exception for all phenomena or hold only for most, many, or some phenomena. For instance, the account of reductive explanation I will present is restricted to the field of biology (first sense of non-universality). In addition, some of the characteristics of reductive explanations I propose are universally applicable (thus, I characterize them as necessary conditions) and yet others are frequent but not ubiquitous. Hence, besides the fact that its scope is limited to biology my account also includes theses that *apply widely* but not universally (second sense of non-universality).

Wimsatt is one of the few philosophers of science who has carved out the second sense of non-universality. As outlined above, he states that adequate meta-theories about science are "as full of *ceteris paribus* qualifiers as [their] subject matter" (2007, 30). According to his view, an adequate philosophy of science should not seek after exceptionless generalizations and analyses in terms of necessary and sufficient conditions since insisting on an "exact, precise, complete, exceptionless description [of scientific practice] can hide important order that is there" (2007, 33). Rather, philosophers of science should account for the "particularity" and "context sensitivity" (2007, 27) of scientific practice.[27] I think Wimsatt draws attention to a crucial point. The diversity of scientific practice forces naturalistic philosophers of science not only to concentrate their analysis on specific scientific fields, but also to come up with generalizations that (at least frequently or sometimes) have exceptions and include *ceteris paribus* clauses and to introduce adequate distinctions that capture crucial differences present in a field (cf. Woodward 2003, 7).

4.2 Balancing Specificity against Generality

On the basis of the above considerations a second criterion of adequacy for an account of reduction can be formulated:

> (Presupposing that the aim is to understand biological practice,) an account of reduction is adequate only if
>
> (2) it captures the *diversity* of the cases of reduction that are present in contemporary biology, but is nevertheless *coherent* and as *universal* as possible.

As this criterion reveals, I do not want to argue for the thesis that philosophy of science should become highly specific or pluralistic, that is, that it should focus on very small areas of scientific research (first sense of non-universality) and exclusively aim at introducing distinctions and developing generalizations that apply only to some phenomena in the considered scientific area (second sense of non-universality). By contrast, I think that philosophers must develop theories that are coherent and general. But they also need to account for the diversity of scientific practice and to be willing to tolerate pluralism without saying what reduction "really" is (Sober 1999, 559). To put it in another way, they are faced with two conflicting demands. On the one hand they must pay attention to the details of scientific practice, for instance by recognizing relevant differences and by not sweeping existing exceptions under the table. On the other hand it is the task of philosophers to

---

[27] Similar reasons prompted Hüttemann and Love (2011) to stress that in their paper on reductive explanations in the biological sciences they are not developing a *theory* of reductive explanation, but rather highlighting different *aspects* of reductive explanations.

abstract from some details of scientific research and to aspire after developing accounts that are coherent and general. Thus, philosophy of science always involves an *act of balancing* specificity against generality – true to the motto "as much generality as you can get and as much specificity as you need".

Alan Love describes this balancing process as a "movement back and forth… on the continuum of abstraction and generality" with the aim of navigating a path between "the twin dangers of losing touch with actual scientific research and becoming a partisan in ongoing explanatory controversy" (2008a, 75). The twofold danger that Love describes here is the following: The more general a philosophical account is and the more it abstracts from the specific empirical data about scientific practice the greater is the risk to lose contact with actual science. In turn, the more particular or non-universal (in both of the two respects distinguished above) a philosophical account is the greater is the danger that it is not philosophically interesting or even loses its status as being "philosophical" at all.

All in all, it seems to me that it can only be determined on a case-by-case basis how exactly to balance specificity against generality and, thus, how much pluralism or non-universality we in fact need. Nevertheless, I think it is important to introduce the above criterion of adequacy and to stress that in developing an account of reduction one should not only seek after exceptionless generalizations but also try to capture the apparent diversity of reductive reasoning practices. To do so requires for example restricting the scope of an account to a certain scientific field (first sense of non-universality) and introducing relevant distinctions and generalizations that apply widely but not universally and not under all conditions (second sense of non-universality).

4.3 The Pluralist's Response

Finally, let me mention a possible objection. A pluralistically minded philosopher of science will probably object that the kind of pluralism (or non-universality) I am concerned with in this section does not go far enough. He could insist that the diversity of scientific practice demands not only that different philosophical accounts are formulated with respect to different scientific fields and that these accounts include also distinctions and *ceteris paribus* generalizations. In addition, the pluralist could argue, there exist several scientific phenomena (e.g. the concept of fitness or of a gene), which require a *plurality of philosophical accounts*. And although there may be inconsistencies between these different accounts, the pluralist could insist that any of the accounts may turn out to be adequate and indispensable to *fully capture* the scientific phenomenon. This thesis is for instance endorsed by Kellert, Longino, and Waters in the introduction to their volume on "Scientific Pluralism" (2006). A slightly different version of scientific pluralism (as a primarily epistemic thesis) is Mitchell's "integrative pluralism" (2003, 2009).[28]

---

[28] Cartwright (1999) and Dupré (1993) have also advocated pluralistic theses with respect to metaphysics.

The thesis that philosophy of science should be pluralistic (or at least allow a plurality of philosophical accounts) is often motivated by the assumption of *pluralism in the sciences* itself (which is also the primary focus of the debate). Accordingly, the volume of Kellert, Longino, and Waters (2006) contains primarily papers which show that a certain area of scientific investigation is presently characterized by an ineliminable multiplicity of approaches. In their introduction they outline:

> The case studies in this book indicate that science provides good evidence that […] some parts of the world (or situations in the world) are *such* that a plurality of accounts or approaches will be necessary for answering all the questions we have about those parts or situations. (2006, xxii)

According to the "pluralistic stance" (2006, xiii) of Kellert, Longino, and Waters, the existing plurality of accounts in the sciences represents an "ineliminable character of scientific inquiry and knowledge (about at least some phenomena)" (2006, xi) and should not be judged as a deficiency in knowledge.[29] When estimating the consequences of this pluralism *in* the sciences the authors allege that it gives rise to pluralism *about* the sciences (i.e. in philosophy of science). For instance, they claim that in the sciences *and* in philosophy of science

> different descriptions and different approaches are sometimes beneficial because some descriptions offer better accounts of some aspects of a complex situation and other descriptions provide better accounts of other aspects (2006, xxiv).

Correspondingly, philosophers of science should refrain from searching for the *one* analysis of a concept that "will enable the pieces to fall into a single representational idiom" (2006, xxv). Since science is a complex phenomenon, the argument of the pluralist runs, several of its elements (e.g. explanation, confirmation, concept of a function, or reduction) cannot be understood by adopting a single philosophical approach.

In this section I do not want to pass a definitive judgment on the plausibility of this strong pluralistic claim. Not surprisingly, scientific pluralism appeals to me. But in order to be a truly convincing thesis – not only about the sciences but also about philosophy of science – the "pluralistic stance" of Kellert, Longino, Waters, and others needs to be further elaborated and specified on basis of examples. Otherwise it remains just too unclear why and under which conditions the incoherency (or even inconsistency) between different philosophical accounts of the same subject is a proper multiplicity that should be sustained and does not represent a deficiency in philosophical knowledge that needs to be eliminated.

However, for my purposes here it is not important whether or not scientific pluralism (as a metaphilosophical claim) can be defended or not. As should have become clear in the

---

[29] These statements reveal that Kellert, Longino, and Waters conceive the kind of pluralism they endorse as an epistemic and not as a metaphysical thesis. Furthermore, they underline that the form of pluralism they advance is not even based on metaphysical assumptions, but rather is empirically motivated (2006, xiii).

forgoing sections, the thesis I am arguing for is weaker than the thesis of scientific pluralism in the style of Kellert, Longino, and Waters. What distinguish our claims are the different kinds of pluralism (or non-universality of philosophical accounts) we advocate. In Section 4.1 I have argued that we need a plurality of accounts with respect to *different* elements in different scientific fields and that these accounts frequently contain theses that exhibit a non-universal character (i.e. apply only widely but not universally or contain *ceteris paribus* clauses). By contrast, Kellert, Longino, and Waters endorse the stronger claim that also with respect to *the same* element of scientific practice there exists a plurality of different philosophical accounts. What is common to both of our pluralistic theses is that they themselves are non-universal. I claim that several (but not all) generalizations in philosophy of science are exception ridden and hold only under certain conditions and the scientific pluralists also argue that several (but not all) scientific phenomena are such that we need a plurality of (possibly mutually inconsistent) philosophical accounts.

5 Philosophy of Science – Descriptive and Normative at Once?

In Section 1 I characterized the account of reduction I will develop in this book as descriptive, naturalistic, and bottom-up. In Section 2 I conceded that this is not the only way to analyze reduction in biology. Alternatively, one could discard the criterion of descriptive adequacy altogether and become engaged in a normative project, which does not seek to understand what reduction in biological practice is, but rather posits what reduction ideally *should* be (first dimension of normativity). In Section 3 I clarified that a descriptive project in philosophy of science must necessarily involve critical-normative elements in order not to collapse into a purely descriptive kind of sociological, historical, or psychological enterprise (second dimension of normativity).

The starting point of this section is the observation that many philosophers of science want to belong to the descriptive, naturalistic, bottom-up layer but, nevertheless, retain their normative aspirations. They stress that the account they propose is "jointly normative and descriptive" (Mitchell 2009, 4; see also Craver 2007a, vii; and Woodward 2003, 7). Or they claim that "[w]ithout being normative [an account]… is not a philosophical account" (Wimsatt 2007, 26) at all. What is crucial to note is that these philosophers do *not* claim that they are involved in a normative project of the kind I delineated in Section 2.1 (first dimension of normativity). That is, they do not abandon the goal of understanding actual scientific practice. Rather, the project they pursue is primarily descriptive and the claim is that this descriptive project involves also normative theses or is "the first step in a normative project" (Craver 2007a, viii). In order to class this kind of project with the ones that I have already outlined, recall Figure I.2 that I presented in Section 2.3:

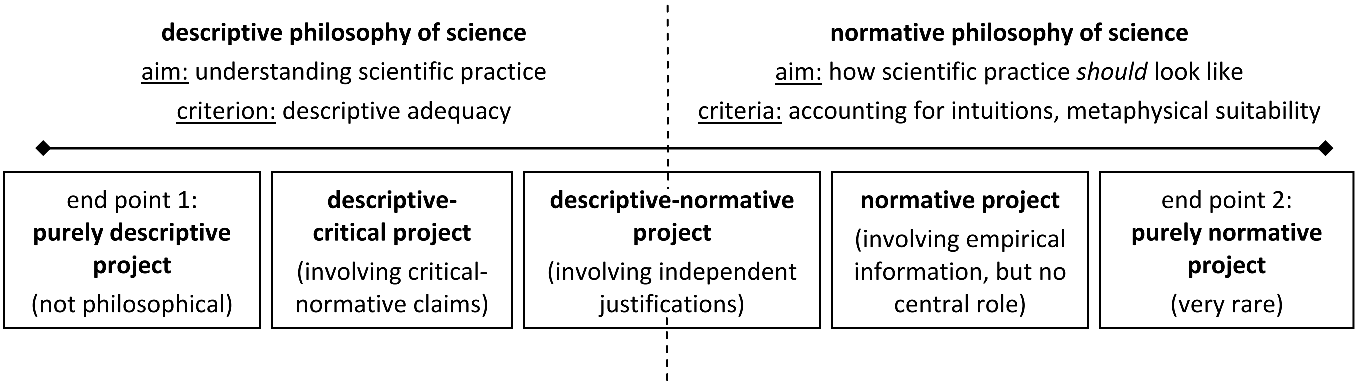

Figure I.2: A Continuum of Different Kinds of Projects in Philosophy of Science

This figure illustrates my initial assumption which shall be validated in this section: accounts like Craver's that are supposed to be "both descriptive and normative" (2007a, vii) are located on the boundary between descriptive and normative projects (represented by the dashed line). That is, on the one hand, projects like Craver's are in a stronger or more extensive sense normative than descriptive-critical projects, because they do not only try to understand scientific practice by employing a critical-normative methodology (second dimension of normativity). In addition, they make independent assertions about how science ideally should be practiced, in particular, about which epistemic norms are independently justified. On the other hand, descriptive-normative projects must be distinguished from normative projects since they do not abandon the aim of understanding actual scientific practice and exclude empirical information from the core of their analysis. In other words, their normative claims (e.g. their independent justification of why certain epistemic norms should apply to science) are not mere philosophical ideals that are detached from real science. Rather, they are "informed" by the results of their descriptive-critical analysis.

The central task of this section is to check whether philosophical projects that claim to be descriptive and normative at once can in fact be classified in this way. In order to assess this we need to examine what it means that the normative part of the project is "informed" by the descriptive part, that is, what relation exists between descriptive and normative claims, what exactly these normative claims are, how this third notion of normativity can be distinguished from the other two dimensions (see also Section 2.2), and whether this descriptive-and-at-once-normative methodology can persuasively be specified.

I begin by analyzing the notion of an (epistemic) norm of science (Section 5.1). This focus is due to the fact that proponents of a descriptive-normative project typically refer to epistemic norms (or virtues) in order to explain what the normative character of their account is. For instance, Craver states that one aim of his account of mechanistic explanation is to "provide a… set of norms by which [mechanistic] explanations should be assessed" (2007a, 111). I specify the notion of an epistemic norm by consulting Water's work on conceptual analysis in philosophy of science (2004). Subsequently, I highlight an important distinction, namely the one between describing epistemic norms that are in fact accepted in science and positing epistemic norms that should apply to science (Section 5.2). I argue that this difference can be interpreted as a difference between the second and the third dimension of normativity and that many philosophical theories are normative according to the second, not to the third dimension (i.e. they are descriptive-critical projects). In Section 5.3 I claim those projects that are descriptive-normative face a dilemma: either they closely connect the descriptive and the normative part of their account and commit an is-ought fallacy, or they avoid this fallacy by detaching the normative part of their account from the descriptive, but cannot make plausible anymore upon which grounds certain epistemic norms that should hold are independently justified. I conclude by pointing out that my own analysis of reductive explanations in biology is normative only according to the second dimension of normativity (Section 5.4). That is, my own account of reduction is a descriptive-critical project, not a descriptive-normative one.

5.1 Epistemic Norms in Science

In his aforementioned paper on concept analysis in philosophy of science, Waters argues that philosophers should “aim to develop an interpretation of concrete scientific knowledge with respect to epistemic virtues such as truth, explanatory power and precision, and predictive success” (2004, 30). He continues by explaining that

> [t]he goal should be to develop accounts of particular sciences that (1) show what epistemic virtues are realized by the sciences (and the extent to which these virtues are realized) *and* (2) reveal a science’s limitations with respect to these and perhaps other epistemic virtues. (2004, 30f)

Thus, Waters’ argumentation seems to be a good place to start trying to understand what these epistemic norms or virtues are, that are supposed to be so relevant to a philosophical theory about science.

Let us have a look at the details of Waters’ arguments. The central question he addresses is how to analyze concepts that are actually in play in science. The special case he is interested in is the concept of a gene. Waters’ main thesis is that concept analysis in philosophy of science is “aimed at something deeper than accounting for linguistic behavior (the use of terms)” (2004, 33f) and “trying to read off the thinking of scientists” (2004, 55) from their actual usage of a concept. But what is this “something deeper”? According to Waters, a scientific concept, like the concept of a gene, should be analyzed in a way that helps us to understand

> how the sciences work (and don’t work) with respect to epistemic virtues that we value… such as the science’s explanatory power and range, the science’s predictive success and failures, the science’s investigative strategies, and the science’s investigative reach (2004, 48).

For instance, Waters argues that an adequate analysis of the concept of a gene needs to be more than a mere description of how scientists themselves think about genes or how different scientists use gene terminology (as it is for example captured in poll-based studies; see e.g. Stotz/Griffiths 2004). Beyond that and more importantly, a philosophical analysis of the gene concept should clarify the relation the concept has to certain epistemic virtues or norms. This means that it should help to clarify the explanatory power and limitations of gene-based explanations and help to account for the investigative utility and biases of gene-centered sciences. To put it another way, according to Waters a philosophical account of a gene should, for example, elucidate why genetic explanations are adequate with respect to some phenomena and inadequate with respect to others and why the search for genes as *the* causes of a phenotypic trait has been successful in some contexts and not in others (e.g. concerning complex diseases; see e.g. Buchanan, et al. 2006).

In sum, Waters' main thesis is that epistemic virtues or norms should take center stage in a philosophical analysis of a certain element of science. But where exactly do these epistemic norms come from? Waters stresses that a scientific concept should not be understood "only with respect to the epistemic values of the practicing scientists" (2004, 50) and that we do not need to "give scientists the final say on what is epistemically valuable about their work" (2004, 53). In other words, Waters claims that a philosophical account cannot and should not merely depict the epistemic norms that are actually endorsed by scientists. Rather, he conceives it to be the task of philosophers to identify those norms that apply to science. And these norms must be the ones that "matter to a broader community" (2004, 49).[30] However, Waters is keen to add that although philosophical judgments are independent, they require "learning a great deal about the content and practice of science" (2004, 55). In order to understand in which sense philosophy of science á la Waters must be normative we need to become aware of an important distinction, which will be introduced in the next section.

5.2 Describing Existing Norms vs. Independently Justifying Norms

Claims about epistemic norms may be said to be normative for two reasons: first, because the *subject* they refer to are epistemic norms or values (and not non-normative events, natural properties, etc.) or, second, because the *claims themselves* are not factual but prescriptive statements (typically indicated by the term 'should'). Correspondingly, a philosophical theory about a certain element of science (e.g. reduction or causation) may be characterized as normative on the grounds of two different features: first, a philosophical

---

[30] To be exact, Waters assumes that epistemic norms that "matter to a broader community" (2004, 49) are of particular importance and that elements of scientific practice should primarily be analyzed in relation to these epistemic norms. Furthermore, Waters emphasizes that there exist no "fixed set of epistemic virtues" (2004, 31), although he admits that truth and empirical success will be valued by anyone interested in science as an epistemic enterprise (2004, 52).

account may be called normative because it describes those epistemic norms that are (at least implicitly) accepted in actual scientific practice. According to this interpretation, the normativity of an account is due to the fact that the subject it describes (i.e. epistemic norms) is normative. An account of this kind contains only *factual statements about norms* (e.g. "The epistemic norm x is acknowledged in current biological research practice"), but not prescriptive statements themselves (e.g. "The epistemic norm y *should* apply to biology"). Second, alternatively, a philosophical account can be characterized as normative because it includes *normative statements* about what epistemic norms *should* hold in science (and not only factual statements about which norms de facto apply to science). One might argue that the first option is a case of normativity, too. However, I am inclined to the view that only philosophical accounts of the second kind, which include normative theses, are genuinely normative. What makes a philosophical theory normative is not that it describes existing norms or values (otherwise, for example, descriptive ethics could not be distinguished from normative ethics), but rather that it entails prescriptive statements (regardless of whether these statements refer to norms or not).

Let us come back to Waters' arguments. The distinction just introduced makes clear that conceptual analysis á la Waters not only involves the mere description of the epistemic norms that are actually accepted in science. Instead, Waters' claim is that philosophers need to make independent judgments, for instance, about the standards according to which the adequacy of gene-based explanations is to be evaluated. In other words, according to him the analysis of a scientific concept must be *normative* because it must contain normative assumptions about which epistemic norms should apply to science, not only factual statements about which norms are in fact employed in science.

It is now clear that the mere description of epistemic norms does not suffice for a philosophical approach to be normative. However, the notion of normativity that underlies accounts that pretend to be descriptive and normative at once still requires further specification. Recall the *three dimensions of normativity* I distinguished in Section 2.2. There I argued that philosophical projects can be normative because they aim at either

(1) developing an account of how science *should* ideally be carried out or how a scientific concept *should* be understood, without taking into account how science actually works (first dimension of normativity; normative projects), or
(2) understanding actual scientific practice, but critically reconstruct the empirical data that scientific practice provides them with (this involves normative claims about which data *should* be included into the account) (second dimension of normativity; descriptive-critical projects), or
(3) understanding actual scientific practice (in particular, the in fact accepted epistemic norms), but also deliver an independent justification of those epistemic norms that *should* be accepted in science (third dimension of normativity; descriptive-normative projects).

Against this background the question arises whether those philosophical accounts that are claimed to be "jointly normative and descriptive" (Mitchell 2009, 4) all belong to the third class of projects (what I call the descriptive-normative projects). To put it briefly, I do not think so. On closer consideration it turns out that several accounts that are said to be descriptive and normative at once in fact belong to the second kind of projects, the descriptive-critical projects. Let me explain this in detail.

In Section 3 I pointed out that the view of philosophy of science as merely depicting scientific practice is a myth. In order to be philosophical each theory about science must involve critical elements such as the choice of paradigmatic and important examples, the explication of assumptions that are only implicitly present in science, and the production of coherence, which involves assessing some statements or actions of scientists as too vague, incorrect, or irrelevant. Philosophers cannot and should not include all information about scientific practice in their account, they need to be selective. And this selection inevitably comprises making normative claims about what counts as paradigmatic, relevant, important, etc. Accordingly, there cannot be a mere description of those epistemic norms that are accepted in scientific practice, as suggested above. Such a description must amount to a critical reconstruction, which involves normative statements about which epistemic norms should be regarded as paradigmatic or as important to science. In a moderate sense these are also claims about which epistemic norms *should* apply to science. However, note that this second dimension of normativity is different from the other two dimensions. Those philosophers who develop a philosophical account in a descriptive-critical fashion aspire to understand scientific practice, not to construct an ideal view of how science should work or how a scientific concept should be understood, that is completely detached from what science in fact looks like (as proponents of normative projects do; first dimension of normativity). Furthermore, descriptive-critical projects are different from descriptive-normative ones (third dimension of normativity) because the former do not aim at making normative claims (unlike the latter), it is just an inevitable component of pursuing the goal of accounting for scientific practice. To put it another way, descriptive-critical accounts critically describe epistemic norms, whereas descriptive-normative accounts add to this critical description an independent justification of certain epistemic norms.[31]

When Mitchell states that her book on "Unsimple Truths" (2009) is "a jointly normative and descriptive project" because the "[a]dequacy of philosophical accounts… are… constrained, but not determined, by the practices and insights of the best contemporary science" (2009, 4) it seems to me that she refers to the second and not the third dimension of normativity. The same is true for Wimsatt who states that "[w]ithout being normative [an account]… is not a philosophical account" and that science should be understood "from the inside while retaining a philosophical perspective" (2007, 26f). To my view, the point they are driving at is that philosophers should not merely passively image

---

[31] What exactly this means will be further elucidated in the next section.

scientific practice, but rather actively and critically reconstruct the practices and insights they find in contemporary science. As I have argued in Section 3, such a critical reconstruction involves making claims about what is important or successful science and what is not. In line with this, Giere states that the goal of a philosophical theory of science is to "explain how science is done and why it is as successful at it is", which involves "making normative claims about how to pursue scientific goals effectively" (1999, 53f).[32] Similarly, Bechtel and Richardson argue that the normative character of naturalistic philosophy amounts to evaluating the strategies scientists use by "identifying contexts where they succeed and where they fail" (2010, 10f). Finally, Waters seems to have a similar point in mind, too, when he claims that

> we (philosophers) should take a critical stance (not necessarily negative, but critical) towards what research scientists publicly claim to be *the* epistemic ideals of scientific knowledge. (2004, 50)

According to Waters, a philosopher should not uncritically take for granted what scientists themselves identify as the proper epistemic norms. Instead, they should come to an "independent judgment" (2004, 54) about which epistemic norms apply to science. However, Waters emphasizes that this does not imply that, for example, "Ruse and Pennock should tell evolutionary biologists how to theorize" (2004, 54). All in all, Waters emphasizes the critical stance and independence of philosophers towards the opinion and self-perception of scientists, but nevertheless requires that philosophy of science must be close to real scientific practice. Accordingly, Waters seems to exclude that the epistemic norms that hold for science can be taken from somewhere other than actual scientific practice or contemporary society.[33]

To conclude, philosophers like Waters, Mitchell, Wimsatt, Giere, Bechtel, and Richardson pursue *descriptive-critical projects* that are normative according to the second dimension of normativity. That is, they concur that philosophical theories about science must transcend mere descriptions and involve normative statements, for example, about what should be regarded as paradigmatic and important examples, which statements of scientists should be ignored as being incorrect or irrelevant, and so on.

5.3 Justifying Epistemic Norms – A Dilemma

Some philosophers of science who claim that their project is both descriptive and normative endorse a stronger thesis. They do not merely assume that their account critically describes epistemic norms that are explicitly or implicitly acknowledged in scientific practice. Rather,

---

[32] Wimsatt makes a similar claim when he demands that a philosophical account of reduction must be "functional" (1974, 700), that is, it must characterize reduction in terms of its functioning in efficiently promoting the aims of science.

[33] In a personal communication Waters has approved of this assumption.

they stress that their project comprises a *justification* of certain epistemic norms, which is independent from their critical description.[34] For instance, in his book "Explaining the Brain" (2007a), Craver not only aims at describing the standards by which neuroscientists evaluate mechanistic explanations. In addition, he addresses the question whether these accepted norms are in fact justified, for example because they produce explanations that are "potentially useful for intervention and control" (2007a, x). Hence, Craver's theory of mechanistic explanations in neuroscience is normative according to the third dimension of normativity. That is, it contains normative statements about what norms of explanation *ideally should* hold in neuroscience. Despite these obvious normative aspirations Craver's approach does not belong to the class of normative projects (first dimension of normativity) because he states that the normative elements in his account are "modeled upon the ideals of neuroscientists rather than those of philosophers and physicists" (2007a, 20).

Craver's statement poses the question of what relation exists between descriptive and normative elements in descriptive-normative accounts. In other words, how is the critical description of epistemic norms that are in fact accepted in science related to the independent justification of those epistemic norms that should apply to science? Those who pursue a descriptive-normative project do not say much about it. For instance, Craver merely admits that the relation between the descriptive and normative project in philosophy of science is "complex" (2007a, viii). Likewise, Woodward claims that they are "interrelated in complicated ways", "mutually inform each other", and, thus, "should be pursued together" (2003, 7f).

The thesis I want to argue for in this section is that Craver, Woodward, and other adherents of a descriptive-normative account encounter a *dilemma* when they try to specify what the relation between their descriptive and their normative theses is. In a nutshell, either they assume a very close connection and commit an is-ought fallacy, or they avoid this fallacy by detaching the normative part of their account from the descriptive. But then they cannot claim anymore that the descriptive part "informs" the normative one and it is left completely unclear on which independent grounds certain epistemic norms are justified.

The first horn of the dilemma is a special version of a well-known objection, which is primarily raised against naturalistic accounts in moral philosophy. Proponents of a descriptive-normative project could assert that there exists a very close connection between, on the one hand, the critical description of those epistemic norms that are in fact accepted in scientific practice and, on the other hand, the normative task of identifying

---

[34] Some authors endorse an even stronger notion of normativity (e.g. Mühlhölzer 2005; Janich 2005). According to their view, an account is normative if it not only includes the *justification* of epistemic norms but also justifies such norms that are *external* to actual science, that is, not (yet) accepted in current scientific practice. Hence, they seem to detach the project of justifying epistemic norms completely from the project of making explicit those norms that are currently accepted in science. In their words, philosophers of science are engaged in normative projects only if they posit epistemic norms for the sciences *from the outside* of science (Janich 2005, 147). It seems to me that this is a too strong notion of the normativity of philosophy of science that is, in particular, improper to accounts that entail normative as well as descriptive elements.

those epistemic norms that should apply to science. They could state that the epistemic norms that *should* hold are simply the ones that are *in fact* accepted in science. Against such a close connection between descriptive and normative theses one could object that this amounts to an *is-ought fallacy*. Philosophers of science cannot simply decide what it is right to do by noting what is actually done. As Bechtel and Richardson put it, "[t]he gap between *is* and *ought* is not this narrow" (2010, 10). If one wants to make claims about which epistemic norms should apply to science, then one cannot simply refer to the fact that particular norms are in fact accepted in science. All in all, the merit of choosing the first horn of the dilemma is that it becomes intelligible what the relation between the descriptive and the normative claims is and on which grounds certain epistemic norms are said to be justified. But the drawback of this choice is that the normative project degenerates to the descriptive project plus a stamping of one's foot and an insistence that the norms actually accepted in science are also the ones that should be accepted.

However, it seems as if philosophers like Craver and Woodward want to try to avoid the is-ought fallacy by emphasizing that they do not simply infer which norms *should* hold from the (critical) description of which norms *in fact* hold. Rather, they argue that their normative claims are solely "informed" by the facts about accepted norms. They seem to argue something like this: although the normative project begins with describing those norms that are accepted in current scientific practice, it clearly transcends this description because it gives reasons for accepting certain norms. For instance, Craver writes that his account is normative not only because it makes the norms of mechanistic explanation that neuroscientists endorse explicit (2007a, x), but also because it *justifies* certain norms (2007a, viii, x). This leads us to the second horn of the dilemma. In order to avoid the is-ought fallacy proponents of a descriptive-normative account must adduce reasons for accepting certain epistemic norms that are independent from and do not amount to the mere fact that they are accepted in current scientific practice. Unfortunately, we are left in the dark about what these independent reasons might be that serve to justify specific epistemic norms. And what is worse, the need for an independent justification conflicts with the assertion of Craver and Woodward that the normative theses are "informed" by or "modeled on" the descriptive ones. It seems as if advocates of a descriptive-normative account can only avoid the is-ought-fallacy if they give up the assumption that the normative part of their project is connected in whatever way to the descriptive part. But this appears to be an unacceptable or, at least, a very problematic step from the perspective of a naturalistically minded philosopher who aims at understanding scientific practice. If the normative part of the project is completely decoupled from the descriptive one the question emerges why the descriptive part is needed at all and what distinguishes descriptive-normative projects (third dimension of normativity) from normative ones (first dimension of normativity).

In conclusion, descriptive-normative projects encounter a serious dilemma. If they connect the descriptive to the normative part of their project they run into an is-ought fallacy, and if they decouple the normative from the descriptive part, the descriptive theses become superfluous and the account turns into a normative project of the kind I delineated in Section 2.1. Thus, my own analysis of reductive explanations in biology is a descriptive-critical project (second dimension of normativity), but not a descriptive-normative project (third dimension of normativity). In the next section I specify what these normative elements in my account of explanatory reduction are.

5.4 The Normative Elements in My Account

First of all, my account of explanatory reduction involves normative claims since it is a critical reconstruction (rather than a mere description) of the central features that reductive explanations in biology possess. Accordingly, my account is normative according to the *second dimension of normativity* which I outlined in Section 2.2 and amplified in Section 3. In short, my account deviates from a mere description of different examples of reductive explanations in biology in three respects: first, the empirical basis it is built upon encompasses only such examples of reductive explanations and of statements of biologists (about reductive explanation and reductionism) that are paradigmatic and important (i.e. instances of successful or failed science). Second, my account makes explicit assumptions about the features that are only explicit in biological practice. Third, in order to develop an account of explanatory reduction that is coherent and unified I need to ponder over which of the heterogeneous and sometimes conflicting statements of biologists I will integrate, and in what way, into my account. This includes identifying some cases as clear-cut, some claims of scientists as correct, and others as too vague, as incorrect, or as irrelevant. Although not all of these critical elements involve normative assumptions, many of them clearly do. For instance, the critical stance involves normative claims about what should be regarded as paradigmatic and important examples of reductive explanations (which implies identifying examples of successful and failed reductive explanations), which statements of scientists about reductive explanation or reductionism should be characterized as correct, appropriate, and relevant, (and thus be incorporated in my account) and which should be dismissed as being too vague, incorrect, or inappropriate.

Another way to frame my project of specifying the central features of reductive explanations is to say that my goal is to specify the *norms of reductive explanation*. These are the standards or criteria according to which explanations are assessed as reductive or non-reductive. It is important not to confound these norms of reductive explanation with the norms of explanation. What determines the reductivity of an explanation is different from what determines its adequacy. In other words, the difference between reductive and non-reductive explanations does not coincide with the difference between adequate and inadequate explanations (see Chapter IV, Section 2.1.1). What makes the task of reconstructing the norms of reductive explanation (rather than the norms of explanation) so difficult is that biologists are far more interested in the latter than in the former. What primarily matters to biologists is whether a proposed explanation of a certain phenomenon is adequate or not. Whether an adequate explanation exhibits a reductive character is not important on its own (see Section 3).

However, although the norms of explanation do not constitute the central target of my analysis, they nevertheless *do* play a role – namely as epistemic norms. To understand this recall what Waters says about the importance of *epistemic norms* to the philosophical interpretation of scientific knowledge (see Section 5.1). His main thesis is that philosophical analyses of scientific concepts should help us to understand which and to what extent epistemic norms are realized by the sciences, and why science sometimes fails to realize these norms (2004, 30f). The examples of epistemic norms Waters invokes are "truth, explanatory power and precision, and predictive success" (2004, 30). Let us apply this idea to the analysis of the concept of reductive explanation. According to Waters, an account of explanatory reduction should clarify the relation reductive explanation has to certain epistemic norms, for example, in which way reductive explanation promotes or hinders explanatory success or prediction. This is one of the issues I address in Chapter V. That is, besides specifying the conditions under which explanations are reductive or non-reductive (i.e. specifying the norms of reductive explanation) I will investigate which consequences follow from the reductive character of explanations for their adequacy. In other words, I will elucidate why reductive explanations are adequate with respect to some phenomena and inadequate with respect to others and what contribution the different features of reductive explanations make to their success or failure. Furthermore, my account of explanatory reduction addresses the related question under which conditions the corresponding reductive strategies or methods for investigating biological phenomena can be applied successfully and under which conditions they fail. This amounts to delineating what many biologists call the merits and "limits of reductionism" (e.g. Ahn, et al. 2006a, 709; Mazzocchi 2008, 10). In order to capture Waters' idea that an account of explanatory reduction should point out in which way reductive explanation promotes or hinders to achieve important epistemic norms (as explanatory success), a third criterion of adequacy needs to be added:[35]

> (Presupposing that the aim is to understand biological practice,) an account of reduction is adequate only if
>
> (3) it clarifies in which sense reductions contribute to or hinder the realization of relevant *epistemic norms*, for instance, explanatory success.

In sum, my analysis of explanatory reduction does not only identify the norms of reductive explanation, it also specifies the relation that reductive explanation exhibits with regard to certain epistemic norms, first of all with regard to explanatory success. My account

---

[35] The requirement that an account of reduction should specify the norms of reductive explanation is captured by the first criterion of adequacy.

is not a descriptive-normative project since it does not include the independent justification of certain epistemic norms. Rather, it critically reconstructs norms that are in fact accepted in contemporary biological practice. That is, my account of explanatory reduction is normative, not according to the third, but according to the *second dimension of normativity*.

6 The Relevance of Philosophy to Science

The way I see it, the most important motivation for developing a critical-descriptive philosophical account about science is that such an account has the best potential to be *useful* to scientific practice (and, perhaps, to the broader society). Actual utility of philosophy to science requires, on the one hand, that philosophy transcends being a mere description of contemporary scientific practice and provides some kind of normative guidance to scientists. On the other hand, a philosophical account can only provide normative guidance and be relevant to scientists if it is close enough to actual scientific practice and if scientists are able to recognize themselves in the philosophical picture. In other words, philosophers should not be afraid to give advice to scientists, but our advice should be "contextual and sensitive to feedback, not a priori pronouncements offered ex cathedra" (Wimsatt 2007, 27). Thus, philosophy of science must at the same time be far enough away from and close enough to actual scientific practice.

With respect to this argumentation the following objection could be raised: Why should philosophy of science aim at being relevant to science (or to the broader society) in the first place? Would it not be too strict to insist that an account in philosophy of science is only adequate if it is judged as being useful by scientists? My response to an objection of this kind is twofold.

First, as the statement by the Nobel Prize winning physicist Richard P. Feynman ("Philosophy of science is about as useful to scientists as ornithology is to birds") shows, philosophers are right about fearing that the quality of their accounts shall be assessed on the basis of whether scientists actually judge these accounts as useful for their work or not. However, I think this implies only that it cannot be the *actual* utility to science that co-determines the adequacy of a philosophical account, but only its *potential* to be useful to science (under the right circumstances). Furthermore, this potential utility cannot be assessed by a minority of perhaps ignorant and prejudiced scientists, but rather must be evaluated in an "ideal conversation" by "well informed and mutually engaged deliberators", as Philip Kitcher envisions it in his "ideal of well-ordered inquiry" (2011). Thus, the following fourth criterion of adequacy for a philosophical account of reduction can be formulated:

> (Presupposing that the aim is to understand biological practice,) an account of reduction is adequate only if
>
> (4) it is *potentially useful* to contemporary biological practice (and to the broader society).

Second, since I want to remain tolerant towards other, purely normative approaches of practicing philosophy of science (see Section 2) I should emphasize that the criterion of potential utility (just as the other three criteria of adequacy) holds only for philosophical accounts that are not purely normative. If a philosopher abandons the aim that his account (at least partially) should capture what goes on in actual scientific practice one cannot require that his account must be potentially useful to science. What is more, it seems to be even misleading that a philosophical account that is completely detached from real scientific practice claims to provide normative guidance, for instance, for how scientists should use certain concepts or how they should reason. Philosophical advice that is based on purely normative accounts would amount to "pronouncements offered *ex cathedra*" (Wimsatt 2007, 27) and, as I think correctly, be perceived as arrogant and as an overestimation of philosophy (e.g. Mühlhölzer 2005, 49). A purely normative account is just too far away from actual scientific practice to be potentially relevant to science.

However, the vast majority of accounts that are developed in philosophy of science and in philosophy of biology, in particular, are not of this purely normative kind. With respect to these accounts I think that potential utility is an important criterion of adequacy. Hence, I agree with Wimsatt that philosophers of science should not only be "theorists of reason", but also "*therapists of reason*" (2007, 27).

7 Interim Conclusion

The primary goal of this chapter was to disclose the meta-philosophical assumptions that lurk behind my analysis of reductive explanation in biology. In other words, this chapter served to explicate and (in part) to justify the aims of and the philosophical methodology by which I develop my account of explanatory reduction. This involved answering questions, such as: Should the goal of developing an account of reduction be to capture actual cases of epistemic reduction or to propose an ideal of reduction that need not be realized in practice? What is the role played by empirical information about actual biological practice? Which kind of empirical information is crucial for analyses of reduction (e.g. information about how biologists actually use the term 'reduction', paradigmatic examples of reductions, biologists' discussions about reductionism, or information about what biologists themselves think reduction is)? How do philosophers deal with this information? How much and which kind of pluralism do we need? In which sense can or should the development of an account of reduction involve normative elements? Should the quality of a model of reduction be judged on the basis of its factual relevance to biological research practice? Answering questions like these does not only clarify the goal and methodology of my own analysis of reductive explanation in biology. Secondarily, it also constitutes a fruitful contribution to the general meta-debate about how philosophy of science can and should be pursued.

In Section 1 I pointed out that the central aim of my analysis is to understand what reduction in current biological research practice is. This goal commits me to accept descriptive adequacy as a criterion of adequacy. In other words, my analysis of reduction must take serious the empirical data about reduction that current biological practice provides us with (e.g. examples of reductive explanations, discussions about reductionism in the biological literature, etc.; Section 1.4). This way of carrying out philosophy has also been called naturalistic since it draws attention to interpreting empirical data about biological practice (Section 1.2). And it has been characterized as bottom-up because it typically starts with the data and proceeds “upward”, that is, towards philosophical theories with broader scope (Section 1.3).

Alternative to such a descriptive project one could pursue a normative project (Section 2.1). Proponents of normative projects in philosophy of science abandon the goal of understanding scientific practice altogether and substitute the criterion of descriptive adequacy by other criteria (e.g. accounting for intuitions, etc.). Descriptive and normative projects are completely different kinds of projects, and I argued there is no need to claim that philosophy of biology *must* be pursued in the descriptive way. In Section 2.2 and 2.3 I refined this picture by distinguishing three dimensions of normativity and corresponding kinds of projects, each of which is located on a continuum ranging from purely descriptive to purely normative projects in philosophy of science: first, *descriptive-critical* projects aim at understanding scientific practice and accept descriptive adequacy as criterion of adequacy. However, they hold that the adequate philosophical methodology is not pure description, but rather critical reconstruction. Second, *descriptive-normative* projects aim at understanding scientific practice (in particular, the in-fact accepted epistemic norms). But on the basis of this understanding they deliver an independent justification of those epistemic norms that should be accepted in science. Third, as already mentioned, *normative* projects aim at developing an account of how science should ideally be carried out or how a scientific concept should be understood. Accordingly, descriptive adequacy is rejected as a criterion of adequacy and empirical information plays no central role in their account.

In Section 3 I further amplified the idea that philosophy of science cannot amount to a pure description of scientific practice, but rather must be carried out in an active, critical-normative way. In other words, philosophers must take up a critical-normative stance on the empirical data scientific practice provides them with. In particular, I argued that philosophers should, first, develop their account on the basis of examples that are paradigmatic and important, second, that they should explicate assumptions that are only implicitly present in scientific practice, and, third, that they should aspire to construct a coherent account.

The latter point gave rise to the question of how this demand of coherency can be spelled out. To put it in another way, how much and which kind of pluralism do we in fact need (Section 4)? I started with distinguishing two senses of non-universality (Section 4.1):

philosophical accounts can be non-universal either in the sense that their area of application is restricted to a certain scientific field or in the sense that the theses they contain only hold for most, many, or some phenomena in a certain field. In Section 4.2 I argued that in developing a philosophical theory one is exposed to two conflicting demands: on the one hand, one should account for the diversity of scientific practice, for instance by recognizing relevant differences and by not sweeping existing exceptions under the table. On the other hand, one should abstract from details and seek to develop accounts that are as coherent and general as possible. Thus, philosophy of science always involves an *act of balancing* specificity against generality – true to the motto "as much generality as you can get and as much specificity as you need". Finally, I countered the objection that this view of philosophy of science leads to accounts that are not pluralistic enough (Section 4.3).

In Section 5 I elaborated on the second kind of projects that I had introduced in Section 2, namely the descriptive-normative projects. The central task of this section was to clarify what it means when philosophers of science claim that their account is jointly descriptive and normative. Several authors link the normativity of their project to the fact that it aims at disclosing certain epistemic norms, like the criteria according to which reductive explanations in biology are assessed as adequate or inadequate. In order to clarify the notion of normativity that is involved here I first illustrated the concept of epistemic norms in science (Section 5.1). Then I revealed an important distinction, namely the one between describing epistemic norms that are in fact accepted in science and positing epistemic norms that shall apply to science (Section 5.2). I argued that several philosophical accounts are normative in the sense that they are critical descriptions of existing norms (second dimension of normativity), not in the sense that they provide independent justifications of epistemic norms (third dimension of normativity). In Section 5.3 I claimed that those philosophers who seek to justify which norms should apply to science face a dilemma: either they commit an is-ought fallacy or they fail to specify the grounds on which their independent justifications are made. I concluded by pointing out that my own analysis of reductive explanations in biology is normative only according to the second dimension of normativity (Section 5.4).

In the last section of this chapter (Section 6) I addressed the question whether philosophy of science should be judged according to its relevance to science (and to the broader community). I argued that it is implausible to regard actual utility as a measure of the quality of a philosophical account. Instead, philosophical accounts should be potentially useful to scientific practice.

In sum, I propose the following four criteria of adequacy for an account of reduction in biology (given that the aim is to understand reduction in contemporary biological practice):

An account of reduction is adequate only if

(1*) (a) it captures the *paradigmatic* and *important* cases of reduction that occur in current biological research practice,

(b) it *explicates* the understanding of reduction that is (often only implicitly) present in biological practice [Section 1 and 3],

(2) it captures the *diversity* of the cases of reduction that are present in contemporary biology, but is nevertheless *coherent* and as *universal* as possible [Section 4],

(3) it clarifies in which sense reductions contribute to or hinder the realization of relevant *epistemic norms*, for instance, explanatory success [Section 5], and

(4) it is *potentially useful* to contemporary biological practice (and to the broader society) [Section 6].

II Drawing Lessons from the Previous Debate

> *"Reduction… is the explanation of a theory or a set of experimental laws established in one area of inquiry, by a theory… formulated for some other domain."*
>
> (Ernest Nagel 1961, 338)

> *"[Nagel's] unitary account of reduction has long dissolved, leaving a polyphonic disunity."*
>
> (William C. Wimsatt/Sahotra Sarkar 2006, 697)

Content

The topic of reduction is not new. It has been addressed at different stages in the history of philosophy (e.g. in the mechanical philosophy of the 16th and 17th century and in the vitalism debate) and it has been among the first issues that were discussed when philosophy of biology emerged as a distinct discipline in the 1960s and 70s. The attempt to apply Ernest Nagel's (1961) classical model of theory reduction to biology denotes the starting point of the debate about reduction in the biological sciences. However, since then much time has passed and the debate has moved forward. This chapter serves to introduce the previous debate about reductionism in the philosophy of biology, which has lasted fifty years so far. But this introduction won't be a mere overview. Rather, I present what I take to be the most crucial lessons one should learn from this debate. In doing so, I introduce and specify important concepts and distinctions and I show the reader the path I will run in the remaining part of this book.

In the course of this chapter it turns out that there are four lessons one should draw from the previous debate: first, before you discuss the correctness of reductionism you should understand what reduction is (Section 1). Second, if you seek to understand what reduction in actual biological practice is, you should focus on epistemic, rather than ontological issues. But you should keep in mind that questions about ontological reduction and about the interrelations between epistemic and ontological issues, nevertheless, are not irrelevant (Section 2). Third, before you discuss reduction(ism) you need to specify which kind of reduction you are talking about – ontological reduction, theory reduction, explanatory reduction, or methodological reduction (the latter three are subtypes of epistemic reduction). With regard to epistemic reduction, reductive explanations are a more promising target of philosophical analysis than reductive methods because they are more constrained and less discipline specific (Section 3). And since, fourth, it is time to move beyond Nagel's inadequate model of theory reduction (Section 4) you should seek to understand reduction in biology by analyzing actual cases of reductive explanations. The latter is exactly what I will do in this book.

1 First Lesson: Understanding Reduction before Disputing About Reductionism

When it comes to the topic of reduction what seems to interest people most is whether one advocates reductionism or whether one belongs to the opposite side, the antireductionists, and on which grounds one tries to defend (anti-)reductionism. Philosophers of science have spent entire books arguing for or against reductionism (see for instance Rosenberg 2006; Bickle 1998; and Dupré 1993) and the papers or chapters of books that are devoted to the defense of certain (anti-)reductionist positions are almost innumerable. Moreover, when it comes to discussions about reduction at conferences many people make you feel the need to "take a stand" by assigning yourself to one of the two parties, reductionism or antireductionism. Talking about reduction seems to inevitably go along with choosing a side – since this is what most people do. But be careful about the side you choose, since either you will be accused of being a "fuzzy thinker" or of being too "simplistic" (Byerly 2003, 336).

In this book I do not want to do what most people do. I will not take a stand on whether reductionism in biology is ultimately correct or not. Rather, my aim is to pursue a different project, which I think is in a certain sense prior to and fruitful for the project of quarreling about the correctness of reductionism. My aim is to understand what this thing called reduction is, about which the reductionists and their opponents make certain assertions. To put it in another way, I seek to develop an *account of reduction*, that is, a detailed analysis of what reduction in contemporary biology is. [1]

From my point of view, the project of understanding reduction is not only different from the project of disputing reductionism; the former is also *prior* to the latter. The reason is that any convincing argumentation in favor of a particular version of (anti-)reductionism or antireductionism must be based on a clear and precise understanding of reduction. You cannot plausibly argue for example that all biological phenomena can be reduced to their underlying physicochemical processes if it remains obscure what exactly it means to reduce biological phenomena to physicochemical processes. Is this a claim about ontological or epistemic reduction? How can the relation of reduction be further specified? And what exactly are the criteria of success for such a reduction? My impression is that in the previous debate too many authors were too often occupied with arguing about the correctness of reductionism without paying sufficient attention to investigating at first what this thing called reduction is they are making assertions about. However, with this I do not claim that in the previous debate the project of understanding what reduction in biology is has been ignored altogether. By contrast, it has been pursued and in some cases yielded fruitful results (e.g. Sarkar 1998, 2005; Hüttemann/ Love 2011), but it has not received the amount of attention it deserves and that would be advantageous for discussions about reductionism. Hence, the first lesson that can be learned from the previous debate is: Before you quarrel about the truth of reductionism, understand what reduction is.

The project of developing an account of reduction in biology involves at least two tasks: first, the identification of the *units of reduction*, that is, the two (token or types of) entities[2] between which the relation of reduction exists, the reduced entity and the reducing entity; second, the specification of the *relation of reduction* itself, that is, an explication of

---

[1] Another way to spell out the difference between the two projects is the following: The project of understanding reduction is concerned with the *interpretation question* "What does it *mean* to say, for example, that biology can be reduced to physics?", whereas the project of disputing reductionism seeks an answer to the *truth question* "Is it true, for example, that biology can be reduced to physics?". This distinction is borrowed from the physicalism debate (e.g. Stoljar 2009).

[2] In this passage I use the term 'entity' not as a purely ontological category (i.e. synonymously with 'object' or 'thing'), but rather in its broader meaning, that is, as a place holder for, on the one hand, objects, properties, facts, etc. in the world and, on the other hand, parts of our scientific knowledge (e.g. explanations, theories, methods, etc.).

what it means to reduce an entity to another.[3] As I will spell out in the following sections the account I develop in this book is an account of explanatory reduction in biology, which is focused on the analysis of individual reductive explanations. Accordingly, the units of reduction I identify are the two parts of an explanation: in reductive explanations the description of the phenomenon or behavior of a system to be explained (explanandum) is reduced to the description of the explanatory relevant factors (explanans) (see also Chapter III, Section 2). Given this, the main part of my analysis will be to specify the relation of reduction by pointing out under which conditions the relation of reduction holds, that is, under which conditions an explanation possesses a reductive character. My main thesis is that these conditions can be identified by revealing *necessary and typical features* that the explanans exhibits (in relation to the explanandum) in case of reductive explanations but not in case of non-reductive explanations. But more on this in Chapter V.

Although I think that this project of understanding reductive explanation in biology is valuable in its own right, it has also the virtue of clarifying discussions about *explanatory reductionism*. We can much better argue about the feasibility, the adequacy and, in general, the scope of explaining biological phenomena reductively if we are very clear about what it means to provide a reductive explanation of biological phenomena.[4] In other words, my analysis of explanatory reduction enables reductionists to provide a more exact account of what they claim, and enables antireductionists to be more specific about what they are denying. Moreover, my account enhances discussions about explanatory reductionism insofar as it clarifies not only what the main features of reductive explanations are, but also what the conditions are under which reductive explanations are inadequate. I will come back to this point in Chapter V.

Despite the importance of the "understanding-reduction project" to the "quarreling-over-reductionism project", it is important to note that the latter amounts to far more than the former. If you want to advocate reductionism in biology you need not only to be clear about what you mean by reduction, you also need to specify the claims you endorse with respect to reduction. That is, you need to point out which *version of reductionism* you argue for. As the bulk of literature reveals, there exist a lot of options here. And since "[e]ach new wrinkle in the reductionist position elicits new responses from antireductionists" (Mitchell 2003, 185) and the other way round, the diversity of (anti-)reductionist positions constantly increases. In order to summarize the most important choices that are available I propose the following list of four criteria according to which different versions of reductionism can be classified:

---

[3] The latter issue has frequently been overlooked so far (see e.g. Brigandt/Love 2008).

[4] However, even if you specify which understanding of reduction your (anti-)reductionist position is based on, an opponent may still challenge your position not by criticizing the arguments you offer in favor of (anti-) reductionism, but rather by questioning the specific account of reduction you build your arguments on. Along these lines Weber states: "the problem for reductionism may be with a particular philosophical *theory* of reduction, not with reductionism itself" (2005, 42).

(1) *in practice* vs. *in principle*: Is reductionism a claim about reductions that are, in fact, accomplishable in contemporary biology (reduction in practice) or is it an in-principle claim about reductions that are supposed to be developed in the future of scientific research, in an ideally complete science, by sufficiently large computers, or by the divine mind (reduction in principle)?

Most philosophers who defend a reductionism with a relatively broad scope (e.g. "All biological phenomena can be explained in physicochemical terms.") do not confine their claim to reductions that are indeed feasible in contemporary science (since this would make their claim false). Rather, they argue that the impracticability of some reductions is simply a reflection of the underdeveloped state of current science or of the limitations of human minds, but does not reflect obstacles in principle (see e.g. Rosenberg's defense of "Darwinian Reductionism" 2006). They claim, for instance, that although not all biological phenomena can be explained in molecular terms yet (in practice), they are explainable in principle, since if we knew everything about the molecules that make up biological objects and about the way the molecules are assembled and interact, we could explain any biological phenomenon reductively. However, in-principle versions of reductionism face the difficulty of how to assess the theoretical or in-principle possibility of reductions (Dupre 1993, 95f; Hoyningen-Huene 2007, 188). This caused some philosophers of biology to take up a deprecatory stance on in-principle reductionist claims (e.g. Wimsatt 2006a, 447).

(2) *eliminative* vs. *retentive*: Does a successful reduction imply the elimination of the reduced entity? Or is reduction retentive in the sense that the reduced entities still play a considerable role in biological practice after being successfully reduced?

Proponents of reductionism are typically keen to stress that they do not endorse eliminativism (e.g. Rosenberg 2006, 54, 84; Wimsatt 2006a, 457; 2007, 168-173).[5] That is, they reject the thesis that a successful reduction results in the elimination of the reduced entities.[6] Wimsatt even calls eliminative reduction a "serious error" (2007, 168) and argues that there is "no evidence for such [interlevel] elimination in the history of science, and there is no reason… to expect it in the future" (Wimsatt/Sarkar 2006, 700). This vehement refusal of eliminativism has also a science-political background: If reductions were eliminative they would call into question the significance of higher-level disciplines, which could have consequences for the distribution of research funds. But this is something that (almost) nobody wants.

---

[5] As I will argue in Chapter III, it is unclear by far whether Rosenberg's Darwinian reductionism is really not eliminative, since what Rosenberg claims is that the phenomena to be explained, which functional biology identifies, are preserved, but not the explanations that functional biology offers (2006, 54, 84).

[6] With respect to elimination there is an important difference between successional and interlevel reduction (see Section 3.4): In case of successional reduction, elimination (i.e. replacement) occurs only if reduction fails. By contrast, interlevel reductions (successful or not) are typically not eliminative. For further details see Wimsatt/Sarkar 2006, 698-700 and Wimsatt 2006a, 457.

(3) *scope*: Does reductionism apply to all biological entities or is it restricted to certain kinds of biological entities (e.g. to a certain biological field)?

Apart from the broad claim that biology can be reduced to molecular biology or all the way down to physics (which is too unspecific to be convincing), there are various reductionist positions available that are restricted in their scope of application. The reductionism debate in philosophy of biology began with the question whether classical genetics can be reduced to molecular biology (e.g. Hull 1972, 1974; Schaffner 1974a, 1993; Kitcher 1984; Waters 1990). In recent decades other biological disciplines were also brought into focus. One dispute concerns for example the putative reductive relation between evolutionary biology and molecular biology (e.g. Beatty 1990; Brandon 1996; Rosenberg 2006), which involves questions as to whether functional explanations can exhibit a reductive character or be reduced to causal-mechanistic explanations and whether evolutionary explanations can be given exclusively on the genetic level (e.g. Wimsatt 1980; Dawkins 1982). Another possibility is to focus on the reduction of developmental biology to molecular biology (Rosenberg 1997, 2006; Laubichler/Wagner 2001; Frost-Arnold 2004).

(4) *locus*: Does reductionism claim that there exists only *one* fundamental level (or "realm" Sarkar 1998, 43) to which the biological entities in question can be reduced? Or does reductionism allow *different* reducing levels (or realms)?

Whereas the previous question concerned the range of the *reduced* entities, this classification criterion relates to the range or, more precisely, the level(s) on which the *reducing* entities are located. There are (at least) two options available: All biological entities in question can be reduced to the entities located on only one, fundamental level (e.g. the level of molecules or genes). Alternatively, a reductionist can claim that different reductions bottom out at different levels (see Chapter V). Borrowing the terminology from Robert Brandon, the former version of reductionism can be called "single-level reductionism" and the latter "multi-level reductionism" (1996, 182).

These four criteria (in practice vs. in principle, eliminative vs. retentive, scope, and locus) provide useful means to clarify the main theses of a particular (anti-)reductionist position and thereby to classify different versions of (anti-)reductionism. Their utility will become apparent in the course of this book when I apply them, for instance, to characterize the kind of reductionism Rosenberg defends (see Chapter III, Section 1.1).

Finally, I should mention that my project of developing an account of reduction is *not completely neutral* with respect to the project of defending (anti-)reductionism. In other words, it is possible that the former yields results which constrain the range of (anti-) reductionist positions one can plausibly advocate. For instance, my analysis will reveal that reductive explanations are not confined to fundamental-level explanations, but that they also include lower-level explanations which do not refer to lowest- or fundamental-level entities such as molecules or genes (see Chapter V). If one presupposes such an understanding of reductive explanations, it would be weird or even incoherent to argue for a version of explanatory reductionism which identifies the level of molecules and genes as the *only* level on which phenomena can be explained reductively ("single-level reductionism" Brandon 1996, 182). In addition, my analysis reveals that reductive explanations are an important element of biological practice (if not the only or the most important element). This fact makes a radical explanatory antireductionism, which denies that biological phenomena can be reductively explained altogether, strongly implausible.

However, my proposal of how to understand reductive explanation in biology leaves many options still open. For example, it does not (at least not directly) prescribe what an adequate scope of (anti-)reductionism may be. What is more, my account of reduction can serve as the basis for in-practice explanatory (anti-)reductionism as well as for in-principle explanatory (anti-)reductionism. At first glance, this may sound odd since my analysis is focused on the critical reconstruction of reductive explanations that can be found *in practice* (and not on reductive explanations that are only imaginable in principle). But an account of what reductive explanation in contemporary biology is can, nevertheless, provide the basis for a version of reductionism that makes claims about whether biological phenomena can in principle (i.e. in the future of biological research) be explained in a reductive manner.[7]

All in all, the first lesson that can be learned from the previous debate is: Before you interfere in the dispute about whether reductionism or antireductionism ultimately is correct, you should aspire to understand what reduction is. This involves identifying the units of reduction as well as specifying the relation of reduction (different kinds of reduction are distinguished in Section 3). Only if you have clarified the notion of reduction can you convincingly argue for a certain version of reductionism (which can be classified according to the four criteria: in practice vs. in principle, eliminative vs. retentive, scope, and locus).

2 Second Lesson: It Is Epistemology that Matters Most

Philosophers who are at home in philosophy of mind are often surprised when they recognize that in the reductionism debate in philosophy of biology *ontological* questions are not the focus of the debate (since in philosophy of mind they are). Instead, *epistemic* questions – for example about the putative reduction of a theory from one field to a theory from a different field, about the adequacy of reductively explaining a higher-level phenomenon in lower-level terms, or about the fruitfulness of applying reductive strategies in certain research fields – are of peculiar interest. In fact, most philosophers of biology quite

---

[7] Granted, this combination of an account of reduction-in-practice with in-principle reductionism is rare and perhaps difficult to balance out. For instance, Rosenberg fails to cope with this balancing act. He argues for in-principle explanatory reductionism but bases his arguments on an understanding of explanatory reduction that is inadequate to biological practice (see Chapter III, Section 1.3). By contrast, most authors who focus on reduction in practice also refrain from advocating in-principle versions of (anti-)reductionism (e.g. Wimsatt 2006a, 2007; Sarkar 1998).

rashly agree on ontological reductionism (in shape of a token physicalism, see Section 2.1) and proceed to the "really" important issues, namely to questions about epistemic reductionism. Their argument typically goes something like that: Vitalism (i.e. the kind of substance dualism that is of particular interest with respect to biology) has proven to be wrong a long time ago. Nowadays, it is obvious that each elephant, blood cell, and enzyme is constituted by nothing but physical objects (i.e. quarks – or whatever the smallest physical particles in fact are). As Rosenberg puts it:

> [S]ubstance dualism about biology… [is] just not a live option. The only biologists who deny physicalism are an assortment of cranks and creationists to whom serious science pays no heed. *We're all physicalists now*. (2006, 4; my emphasis)

Hence, the case of ontological reductionism is closed since everybody accepts physicalism as default position. Consequently, philosophers of biology turn to the issues that are of "real" importance and examine what follows from ontological reductionism for questions about epistemic reductionism.

As the title of this section indicates, I agree that the epistemic questions concerning reduction are the ones that matter most to biology (see Section 2.3). However, I also think that philosophers of biology would be wise not to dismiss the ontological questions in such a hasty manner. There is more to be said about these issues than "We're all ontological reductionists. Case closed". Furthermore, the disputes about reductionism show that epistemic questions often are entangled with and get mixed up with ontological ones. Not uncommonly, proponents of reductionism derive their arguments for epistemic reductionism from ontological reductionism without putting their cards on the table (e.g. Rosenberg 2006). Thus, even if the reductionism debate in philosophy of biology correctly is primarily concerned with epistemic issues, philosophers need to pay attention to the interrelations that exist between questions of ontological and epistemic reduction, too.

On these grounds, in what follows I clarify the notion of ontological and epistemic reduction (Section 2.1) and, then, I try to shed light on the connections that exist between these two issues (Section 2.2). I conclude by showing why I nevertheless think that epistemic issues are more relevant to biology than ontological ones (Section 2.3).

2.1 Ontological and Epistemic Reduction

In short, ontological reduction is a specific relation that exists between *entities*[8] *in the world* (i.e. between objects, properties, events, facts, processes, states, or whatever you think is the basic inventory of the world), whereas epistemic reduction is a relation that holds between certain parts of our *knowledge* about this world (e.g. between theories from

---

[8] The term 'entity' is used here in its wider meaning, that is, it is not confined to objects or things but refers to all kinds of "what there is".

different scientific fields, between our description of the phenomenon to be explained and the description of the explanatory relevant factors in a reductive explanation, etc.).

This difference between what there is in the world (ontology) and what our representations or knowledge about the world are like (epistemology) is crucial – not only for the issue of reduction. What holds for ontology need not also apply to epistemology. You can have ontological reduction without, at the same time, having epistemic reduction.[9] This difference enabled philosophers of biology to be ontological reductionists and, nevertheless, to form the "antireductionist consensus" (Waters 1990), that is, to endorse epistemic *anti*reductionism. Although there exist some philosophers who challenge this combination of ontological reductionism and epistemic *anti*reductionism as an "unstable equilibrium" (Rosenberg 2006, 7; see also Rosenberg/Kaplan 2005), the majority of philosophers think that ontological and epistemological issues fall apart.[10] Before we examine the exact relation between ontological and epistemic reduction let us first dwell on what is meant by ontological reduction (Section 2.1.1) and by epistemic reduction (Section 2.1.2).

2.1.1 Specifying the Notion of Ontological Reduction(ism)

As I have mentioned before, in philosophy of biology ontological reductionism is widely taken for granted, but not really a subject of discussion. It seems to me that, although the focus on epistemic issues is appropriate (see Section 2.3), the debate would yet benefit from becoming engaged in ontological disputes. Avowed statements such as "We're all physicalists now" (Rosenberg 2006, 4) are not satisfactory if it is not further specified what is meant by physicalism and in which sense this physicalism is an ontological-reductionist account. At this point there is much (but not all) to be learned from philosophy of mind.

However, to be fair, it is not the case that philosophers of biology remain completely silent about the kind of ontological reductionism they take for granted. In order to get an overview, have a look at the following collection of quotations:

> [T]here are no major figures in contemporary biology who dispute the claim that each biological event, state, or process is a complex physical event, state, process. (Kitcher 1984, 369)

> Reductionism… often has the connotation that biological entities are '*nothing but*' aggregates of physicochemical entities; this sort of approach can be termed ontological reductionism of a strong form. (Schaffner 1993, 413; my emphasis)

> We live in one world, not many worlds. Further, the material from which all the entities in the world are built is ultimately *one kind of 'stuff'*, that is, matter [i.e. material composition assumption]. (Mitchell 2003, 181; my emphasis)

---

[9] Of course, this depends on the notions of ontological reduction and of epistemic reduction involved. See also Section 2.2.

[10] This is not to say that there exist no close connections between ontological and epistemic issues concerning reduction (see Section 2.2).

> Ontological reduction is the idea that each *particular* biological system (e.g., an organism) is *constituted* by *nothing but* molecules and their interactions. (Brigandt/Love 2008, Section 1; my emphasis)

> The metaphysical thesis that reductionists advance… is physicalism, the thesis that all facts, including all functional biological facts, are *fixed by* the physical and chemical facts; there are no non-physical events, states, or processes, and so biological events, states, and processes are '*nothing but*' physical ones. (Rosenberg 2006, 25; my emphasis)

> The physicalism… nobody is going to deny… is not metaphysical materialism, but what philosophers call the *supervenience* of all phenomena on physical foundations. That is, the properties and relations of the components are taken to *fix* the states of the whole, but the behavior of the whole may *not* be effectively *derivable* from properties of the components. (Byerly 2003, 337; my emphasis)

These citations are far from being equivalent. However, each of them provides important insights into how the notion of ontological reductionism is understood in contemporary philosophy of biology. My central goal in this section is to specify the notion of ontological reduction that is implicit in the above quotations by consulting some important distinctions and assumptions that were introduced in philosophy of mind. That is, I aim at clarifying what philosophers of biology mean when they talk about ontological reductionism, but I do not want to argue that is the way the term 'ontological reduction' *should* be understood. The latter is important to notice since a major result of my analysis will be that the concept of ontological reduction that is accepted in philosophy of biology differs in several respects from how this concept is understood in philosophy of mind.

To put it in a nutshell, my analysis will yield the following results: first, the kind of ontological reductionism that is taken for granted in philosophy of biology is *token physicalism* (or materialism)[11], whereas in wide areas of philosophy of mind the term 'reductionism' is still confined to type or property physicalism. Second, despite this different usage of the term 'reduction(ism)' philosophers of biology agree with philosophers of mind that type or property reduction (i.e. identification) is deeply problematic and that the relation that exists between mental or biological properties and physicochemical properties is one of *supervenience*. Third, when token physicalism is considered, the relation of reduction between tokens (e.g. between a particular blood cell and a particular bulk of molecules or quarks) can either be understood in a strong sense, that is, as an *identity* relation (typically expressed by a nothing-but claim)[12], or in a weaker sense, for instance, as

---

[11] Although there are important historical differences between materialism and physicalism and although one could possibly also point out some systematic differences (as e.g. Dupré 1993, 90-94 does), I will join most disputants and use these two concepts interchangeably.

[12] In a few cases *nothing-but claims* can also be found in contexts where epistemic issues are concerned. For instance, Wimsatt refers with the label "nothing-but-ism" to the "attempt to deny or discredit the explanatory potency of upper level entities, phenomena, and causes" (2006a, 446), that is, to explanatory reductionism. However, it is far more common to interpret statements like "organisms are nothing but bags of genes" (van Regenmortel 2004, 148) as expressing an *ontological* and not an epistemic claim.

a relation of *constitution* or localization. In what follows I will amplify these claims. However, because an extensive debate and an immense bulk of literature in philosophy of mind stand behind each of these claims, in what follows I will only introduce those details that are necessary for understanding the notion of ontological reduction(ism) that is accepted in philosophy of biology.

Type versus Token Reduction

Ontological reductionism can either be a claim about type reduction or about token reduction, depending on what one identifies as the units of ontological reduction (i.e. as the relata between which the relation of ontological reduction holds). *Type reduction* is a relation between different types of entities in the world (e.g. between properties, types of objects, types of processes, etc.)[13], whereas *token reduction* is a relation between particular entities (i.e. instantiations of properties, particular objects, particular processes, etc.). An example of the former would be the reduction of the type 'classical gene' to the molecular type 'DNA strand with certain other features [to be specified]'.[14] By contrast, the reduction of a particular gene, say, for red eye color, of an individual *Drosophila* fly to a particular DNA strand with a specific nucleotide sequence would be a case of token reduction.

As the above quotations clearly show, the kind of ontological reductionism that is presupposed in philosophy of biology is *token physicalism* and not type or property physicalism (which is also known as "type identity theory" or just as "identity physicalism")[15]. That is, the assumption is that each particular biological entity is nothing but a physical entity since it can be reduced to a particular assemblage of physical (or physicochemical) entities. The assumption is *not* that biological properties or types can be reduced to physical properties or types in the sense that the former can be shown to be identical to the latter.[16] This is most evident in the quotes from Kitcher, Brigandt, and Love:

> [T]here are no major figures in contemporary biology who dispute the claim that *each* biological event, state, or process *is* a complex physical event, state, process. (Kitcher 1984, 369, my emphasis)

---

[13] I am aware of the fact that some authors identify types with properties. However, I endorse a broader notion of type here, according to which types can but need not be properties (see Wetzel 2011).

[14] With respect to the problems that occur when one tries to define what a gene is solely in molecular terms (and with respect to possible solutions of these problems) see, for instance, Schaffner 1993, 446-451; Waters 1994; and Weber 2005, Chapter 7.

[15] Early proponents of the type identity theory of mind were for instance Place (1956), Feigl (1958), and Smart (1959).

[16] This is why it is possible to endorse physicalism (namely token physicalism) and, at the same time, to claim that biological systems can possess emergent properties (see van Regenmortel 2004, 146). At first sight, this seems to be a contradiction since the concept of emergence is frequently assumed to be complementary to the concept of reduction (e.g. Kim 1999). However, emergence is typically only opposed to type or property reduction, not to token reduction. Thus, it is possible to accept that biological systems exhibit emergent properties and, nevertheless, to insist that these concrete biological systems exist exclusively of physical entities. For details concerning the relation between emergence and reduction see, for instance, Bedau/Humphreys 2008.

> Ontological reduction is the idea that *each particular* biological system (e.g., an organism) *is* constituted by *nothing but* molecules and their interactions. (Brigandt/Love 2008, Section 1; my emphasis)

What might be surprising from the perspective of philosophy of biology is that most philosophers of mind *confine* the notion of ontological reductionism to *type* or *property physicalism* (e.g. Baker 2009; Kim 2005, 34; Beckermann 1992). In other words, in philosophy of mind ontological reduction is taken to be an identity relation between two different properties of the same system (Hüttemann 2004)[17], namely between mental and physical properties. For example, the desire for eating ice cream is assumed to be nothing more than (i.e. identical to) a specific pattern of action potentials in a certain brain region and can in this sense be reduced to it. Even if philosophers of mind consider different "models of reduction" (Kim 2006, 275; 2008, 96), for instance, bridge-law reduction, identity reduction, and functional reduction, they conceive reduction still as being a *relation between properties* (i.e. between types and not between tokens).

This common restriction of the notion of ontological reduction to the notion of type or property identity is largely due to the great influence Nagel's (1961) classical model of theory reduction (see Section 3.1 and 4) had and still has in philosophy of mind.[18] According to Nagel, cases of heterogeneous reduction require the existence of bridge principles (also known as "bridge laws"), which connect the vocabulary of the reduced theory to that of the reducing theory. Since bridge principles were commonly conceived as expressing identities between types, theory reduction was classically supposed to, ontologically, require type reductions (i.e. type identities).

The subsequent debate in the 1960s to 1980s was characterized by the discussion of many serious problems type physicalism faces (e.g. the problem that many mental and biological types are *multiply realized* on the physicochemical level; see Sections 3.1 and 4). In order to meet these challenges, Hilary Putnam (1975), Jerry Fodor (1974), and Richard Boyd (1980) developed an alternative version of physicalism, which was named *non-reductive physicalism* (or materialism) since it abandons the strong claim that mental properties can be identified with and, thus, be reduced to physical properties. Non-reductive physicalism offers the opportunity to adhere to physicalism but yet to acknowledge the distinctness of mental and physical properties. In these days, a bewildering variety of versions of non-reductive physicalism is advocated (e.g. Pereboom/Kornblith 1990; Pereboom 2002; Baker

---

[17] This is another crucial difference between debates about reduction in philosophy of mind and in philosophy of biology: whereas the former is concerned with the relation between two properties of the same system (i.e. mental and physical properties), the latter discusses the relation between two properties of different systems, namely how properties of wholes are related to properties of parts. The latter kind of reduction is also called micro-reduction. For further details see Hüttemann 2004, 10, 26f, 59.

[18] This situation has only slowly begun to change. For instance, only recently Carl Gillett characterized Kim's abandonment of Nagel's model (he reviewed it as an "inappropriate model of reduction… [that] largely appears beside the point for issues of real philosophical significance"; Kim 1999, 13) as a "seismic shift" that is "slowly impacting a range of debates and issues" (Gillett 2007, 193).

2009; Yoo 2008; Wilson 2010; etc.) and although non-reductive physicalism encounters many objections (see, in particular, Kim 1989; 1993, Chapter 14 and 17; 2005, Chapter 2) it represents the current default stance in philosophy of mind.

In sum, philosophers of mind restrict the concept of ontological reduction to type or property identity and they may have good reasons to do so.[19] By contrast, in philosophy of biology everybody agrees that type or property physicalism is inadequate. This is why the kind of ontological reductionism that "everybody in the debate is going to accept" is spelled out as token physicalism. Hence, philosophers of biology employ a broader notion of ontological reductionism than philosophers of mind, namely one that encompasses not only type but also token physicalism.[20]

What instead of Type Identity?

What is interesting with respect to further clarifying the notion of ontological reduction is that there seem to exist two options for a non-reductive physicalist: either he follows Fodor (1974) and argues for *token physicalism* or he stays with types and looks for an alternative relation between mental or biological and physical properties. (Of course, one can also take both options simultaneously.) The goal is to find a relation that is not as tight as reduction (i.e. identity) but tight enough so that the resulting position still counts as or is compatible with physicalism.[21] The relation of supervenience allowed for a combination of the autonomy of higher-level properties with physicalism. *Supervenience physicalism* is intended to capture the minimal or core commitment of physicalism and therefore is also called "minimal physicalism" (Lewis 1983). The notion of supervenience was introduced by Donald Davidson (1970, 214) into contemporary philosophy of mind and can roughly be formulated as following: "A set of properties *A* [e.g. mental properties] supervenes upon another set *B* [e.g. physical properties] just in case no two things can differ with respect to *A*-properties without also differing with respect to their *B*-properties" (McLaughlin/Bennett 2011). There are many different ways available in which the notion of supervenience can be further specified (for an overview see, e.g., Mclaughlin/Bennett 2011).

Whatever these specifications amount to, what interests me at this point, is that there seems to exists an important difference between the two options outlined above, namely *token physicalism* and *supervenience physicalism*[22]. As I have already indicated,

---

[19] For instance, a philosopher of mind could argue that a concept of reduction that is widened in such a way that it applies also to identity relations between tokens is uninteresting or useless because it is compatible with so many other positions (e.g. with property dualism, too).

[20] One might object that such a broad concept of ontological reductionism renders almost all positions (ontologically) reductionistic, but this does not change the fact that the notion is actually applied in this way.

[21] The fact that non-reductive physicalism tries to combine physicalism with a commitment to the autonomy of higher-level properties (and thereby, allegedly, becomes unstable) is a popular point of attack. See Kim 1989; 1993, Chapter 14 and 17; 2005, Chapter 2. For replies to this challenge see, for instance, Yoo 2008 and Wilson 2010.

[22] One could argue that, strictly speaking, supervenience physicalism is also a version of type physicalism since it is a claim about the relation between types or properties (and not tokens). However, in philosophy of mind a

philosophers of mind label both positions 'non-reductive' because both of them reject type or property identity. By contrast, the majority of philosophers of biology broaden the notion of ontological reductionism in a way that it includes token physicalism as well. This gives rise to the question whether even supervenience physicalism alone suffices to specify the kind of ontological reductionism that any philosopher of biology accepts. The quotation of Byerly seems to suggest this:

> The physicalism… nobody is going to deny… is not metaphysical materialism, but what philosophers call the *supervenience* of all phenomena on physical foundations. That is, the properties and relations of the components are taken to *fix* the states of the whole, but the behavior of the whole may *not* be effectively *derivable* from properties of the components. (2003, 337; my emphasis)

However, if we carefully consider this quote we will recognize that Byerly only claims that supervenience physicalism is the kind of physicalism, but not the kind of ontological reductionism, everybody accepts. This is compatible with the thesis that supervenience physicalism alone is insufficient to be referred to as reductionism. Other statements of philosophers of biology support the assumption that in order to turn supervenience physicalism into an ontological reductionism one needs to add (at least) token physicalism.[23] Recall for instance the quote of Rosenberg:

> The metaphysical thesis that reductionists advance… is physicalism, the thesis that all facts, including all functional biological facts, are *fixed by* the physical and chemical facts; there are no non-physical events, states, or processes, and so biological events, states, and processes are '*nothing but*' physical ones. (Rosenberg 2006, 25; my emphasis)

All in all, you have to draw the line between reductionist and non-reductionist positions somewhere. In philosophy of biology this line is drawn between, on the one hand, type or property physicalism and token physicalism (which are called reductionist) and, on the other hand, supervenience physicalism (which is, on its own, referred to as non-reductionist).

Identity, Constitution, or What?

Another question that is relevant in this context is the question of what *kind of relation* must exist between mental or biological tokens/types and physicochemical tokens/types in order that this relation counts as a reduction.

The answer that has dominated the debate in philosophy of mind thus far is that the relation of ontological reduction is a relation of *identity*. The popularity of this answer traces back to the great influence that Nagel's model of reduction had on the debate and to the

---

narrower notion of type physicalism, which is confined to type *identity* physicalism, has become prevalent. Here I adopt this common understanding of the term 'type physicalism'. According to it, supervenience physicalism is *not* a version of type physicalism.

[23] However, a few philosophers of biology even argue that the assumption of supervenience is sufficient to call something a (weak) reductionist position (see, e.g., Dupré 1993, 94-99).

fact that Nagel's demand for bridge statements (in case of heterogeneous reduction) was interpreted as a demand for identity relations between properties. In recent years critical voices can be heard that question whether there in fact exist relations of identity between the mental and the physical realm, for instance between mental causal powers and microphysical causal powers (Pereboom 2002). A promising alternative seems to be to appeal to the relation of *constitution*. Several metaphysicians have argued for the thesis that the relation of constitution is distinct from the relation of identity (e.g. Baker 1997; 2000, Chapter 2 and 7; 2007, Chapter 2, 8, and 9; Johnston 1992). However, in the philosophy of mind most authors critical with respect to identity regard the relation of constitution as an *alternative* to reduction, which they take to be confined to the relation of identity. This is why, for example, Pereboom calls his position "Robust *Non-reductive* Materialism"[24] (Pereboom 2002; my emphasis; see also Pereboom/Kornblith 1990).[25]

As I have already pointed out, in philosophy of biology the term 'ontological reduction' is understood in a broader manner, that is, as applying to relations between particulars (i.e. tokens), too. This gives rise to the question of which relation exists between biological tokens and, for instance, molecular or physicochemical tokens and which of these relations can be adequately called a reduction. For instance, how can the relation between a concrete red blood cell and a particular assemblage of molecules (predominantly, hemoglobin) organized and interacting with each other in a certain way be further characterized? Is the blood cell *identical* to the assemblage of molecules (organized and interacting in a specific way)? Or is it more adequate to regard the blood cell as being *constituted* by the assemblage of molecules? On which grounds can one distinguish constitution from identity? And is exclusively the relation of identity a *reductive* relation, or can the relation of constitution be treated as reductive, too?

Expectedly, the majority of philosophers of biology who discuss the issue of reduction(ism) do not address these ontological questions at all. The citations from which we started and that express different views about ontological reductionism do not provide clear answers to these questions. Many authors link the notion of ontological reduction closely to the thesis that particular biological entities are "*nothing but*" specific assemblages of lower-level entities (see, in particular, Schaffner 1993, 413; Brigandt/Love 2008, Section 1; Rosenberg 2006, 25). The phrase 'nothing but' or 'nothing over and above' is frequently interpreted as expressing the relation of identity. However, one need not interpret this phrase in this way. It might also be interpreted as expressing the fact that for instance a particular cell is *constituted* by nothing over and above molecules that are organized and that interact with each other in a certain manner. One of the few philosophers of biology who say a bit more about these ontological issues is Wimsatt. In his recent work he explicitly

---

[24] Similarly, Baker (2009) defends a "property-constitution" version of non-reductive physicalism.

[25] Only a few philosophers of mind depart from this and include cases of "compositional reduction" (Gillett 2007) or "microreduction" (Hüttemann 2004) in the class of genuine ontological reductions.

denies Schaffner's claim that identities are required for successful reductions (1967). Wimsatt argues that, instead, "localizations" (2006a, 456; see also Bechtel/Richardson 2010, xxxii) suffice, too, because they "preserve all relevant spatiotemporal properties of identities" (Wimsatt/Sarkar 2006, 700). He states that localization is distinct from identity, more precisely that localizations are "logically weaker than identities" (2006a, 456). However, Wimsatt's few remarks do not help much in answering the ontological questions raised above because, on the one hand, Wimsatt is more concerned with localization and identification as *research strategies* (i.e. with methodological reduction), rather than with localization and identity as ontological relations between entities from different levels. On the other hand, it is difficult to pin down what exactly the suggested difference between relations of identity and relations of localization is, and how the latter relate to relations of constitution.

To conclude, it seems as if in philosophy of biology there is much work left over with respect to specifying the kind of relation that characterizes cases of ontological token reduction. In Chapter V I will argue that part-whole relations (more precisely, relations between biological systems and their genuine parts) are central for determining the reductivity of explanations. This suggests that relations between wholes (e.g. a particular thin muscle filament) and sets of their organized and interacting genuine parts (e.g. the set of certain actin molecules, which are twisted into a helix configuration, tropomyosin molecules, which cover the myosin-binding sites on the actin molecules when muscle fibres are relaxed, and troponin molecules, which are attached to the tropomyosin molecules and allosterically modulates them) are important cases of ontological reduction. However, this leaves unanswered the question whether the relation between a concrete thin muscle filament and a particular set of organized and interacting actin, tropomyosin, and troponin molecules is a relation of identity, a relation of constitution, or even a completely different relation.

To sum up the results of the previous sections, the kind of ontological reductionism philosophers of biology typically take for granted is token physicalism (whereas supervenience physicalism alone is conceived as being non-reductionistic). Accordingly, their notion of ontological reduction varies from one that is accepted in philosophy of mind in an important respect: it is not restricted to a relation between types or properties. What is more, in philosophy of biology it is left open whether the relation of reduction must be one of identity or can be specified as a relation of constitution (or localization) as well or instead (but it remains unclear, too, what the difference between these two relations is).

2.1.2 Specifying the Notion of Epistemic Reduction

Whereas ontological reduction is a relation that exists between entities in the world, epistemic reduction is a relation between parts of our *knowledge* or *representations* of these entities and relations in the world. To put it another way, in the case of epistemic reduction the units or relata of reduction are *epistemic entities* like descriptions, theories, models, explanations, law statements, methods, etc. Because of this variety of epistemic entities that play (or were supposed to play) a role in biological practice, different candidates for the relata of epistemic reduction are available. Depending on the epistemic entity one regards as being most important for reduction in biology, one gets a certain kind of epistemic reduction(ism). To distinguish the different types of epistemic reduction in detail is the task I will approach in Section 3. At this point, a brief overview shall suffice.

For a long time the debate about epistemic reductionism in philosophy of biology was centered on *theories* and *law statements* as the relevant epistemic units of reduction and the relation of reduction was characterized in a formal manner, that is, as a relation of logical derivation. This fact was due to the great influence Nagel's classical model of theory reduction (1961) had on the debate (see Section 4). In addition, at the beginning the debate focused on quite *global* cases of epistemic reduction like the reduction of entire biology to physics (or, at least, to molecular biology) and the reduction of a whole branch of biology to another, for instance, the reduction of classical genetics to molecular biology as the "test of reductionism" (Rosenberg 1985, 90) *par excellence*. In this spirit, for example, Francis Crick made his famous claim that "the ultimate aim of the modern movement in biology is to explain all biology in terms of physics and chemistry" (1966, 10). More than a few biologists still have this kind of global theory reduction in mind when they reject "epistemic reductionism" (e.g. Mayr 1988, 475) or when they stress that "epistemic reduction" is *not* what they are talking about (e.g. Fang/Casadevall 2011, 1401). Furthermore, this was also the way the notion of epistemic reduction(ism) was introduced in philosophy of biology (Ayala 1974, ix).

However, to focus on theories, formal issues, and global cases of epistemic reduction is not the only route that has been taken so far and, as I will argue in Sections 3 and 4, it is not the most promising way to try to understand what epistemic reduction in biology really is. Alternative epistemic entities that have been considered as being important for reduction are *explanations* and *methods* (or heuristics or research strategies). According to this, in Section 3 I distinguish three kinds of epistemic reduction(ism): theory reduction, explanatory reduction, and methodological reduction. Let us now turn to the relation that exists between ontological and epistemic reduction.

2.2 Relations between Ontological and Epistemic Issues

At the beginning of Section 2.1 I had already put my cards on the table. I agree with most philosophers of biology that in the debate about reduction it is important to clearly distinguish ontological from epistemic issues (Sarkar 2005, 106; 1992, 169; Mitchell 2003, 181; Brigandt/Love 2008, Section 1; Schaffner 1993, 515), yet to recognize the relations that hold between ontological and epistemic reduction. But what exactly are these relations?

2.2.1 How Tight Is the Connection?

Consider, at first, the two most radical (and most controversial) answers to this question, namely that you cannot have ontological reduction without at the same time having explanatory reduction (this is what Rosenberg claims), respectively, that ontological and explanatory reduction fall together (this is what Brandon claims). Although Brandon acknowledges the difference between methodological, ontological, and explanatory reductionism, he decides that it "would be tedious to continually make this threefold distinction" and that the "very tight connection between the explanatory and ontological versions" (1996, 180) allows him to lump together ontological and explanatory reduction(ism) under the single label 'ontological reduction(ism)'. Rosenberg assumes a tight connection between ontological and explanatory issues, too. But he does not go as far as to override the difference between ontological and explanatory reduction(ism). Rather, he claims that ontological reductionism strongly enforces explanatory reductionism since "the combination of physicalism and [explanatory] antireductionism… looks like an 'unstable equilibrium'" (2006, 7) or is an "untenable dualism" (2006, 1).

In my view, both authors derive their claims about the very tight connection between ontological and explanatory reduction from a certain view about scientific explanation. Brandon argues only briefly that this tight connection follows from "the causal/mechanical model of explanation" which alleges "a tight link between one's ontological view, one's view about the fundamental entities and processes that make up the world, and one's view of proper explanation" (1996, 180f). This statement is not of much help since it does not spell out why a proponent of the causal-mechanical view of explanation should be committed to the claim that a phenomenon can *only* be explained by reference to "fundamental entities and processes" – irrespective of what the question is that the explanation is supposed to answer.

Rosenberg's defense of "Darwinian reductionism" (2006) reveals more insights, even if his argumentation is unconvincing at several points. In short, according to Rosenberg biological explanations are adequate only if

(1) they refer to a *law* (in case of biology, the only law we have is the chemical "principle of natural selection (PNS)"; Rosenberg 2006, Chapter 4 and 6; Rosenberg/Kaplan 2005) and
(2) they are *maximally complete* (2006, 4, 12, 14, 26), that is, all the causal chains described in the explanans are filled in by molecular details (2006, 46).

In Rosenberg's words, only proximate (i.e. molecular) why-necessary explanations that refer to PNS are adequate explanations in biology (2006, 42-47).[26] What interests me at this point is Rosenberg's second assumption that explanations in biology are adequate only if they

---

[26] The details of Rosenberg's view about reduction and explanation will be fully elucidated and critically examined in Chapter III and IV.

provide a maximally complete description of the causal process that underlies the explanandum phenomenon. Since Rosenberg understands the notion 'being maximally complete' as 'being maximally complete *on the molecular level*', the requirement of maximal completeness inevitably leads to *explanatory reductionism*, that is, to the claim that

> explanations in functional biology need to be corrected, completed, or otherwise made more adequate by explanations in terms of molecular biology (2006, 26).

But why should we adopt the assumption that explanations are adequate only if they are maximally complete on the molecular level in the first place? This is the point where Rosenberg brings ontological reductionism into play. He argues that "[o]nly someone who denied the thesis of physicalism… could deny the causal relevance of… macromolecular processes" (2006, 36). In other words, the molecular processes underlying a higher-level phenomenon are *always* causally relevant. Furthermore, Rosenberg emphasizes that what is causally relevant must also be explanatorily relevant and, thus, be included in the explanation of the higher-level phenomenon. For Rosenberg there can be no deviation of explanatory relevance from causal relevance since he adopts a "non-erotetic approach to explanation" (2006, 44), which rejects the view that what is explanatorily relevant depends also on pragmatic factors like the question being asked or the research interests of biologists. Rosenberg maintains that such a view of explanation results in subjectivism (2006, 35f, 44, 179f). To sum up, ontological reductionism enforces explanatory reductionism – but only if one adds some very specific and highly controversial assumptions about causal relevance and about the determinants of the adequacy of explanations. In Chapter III and IV I will further examine Rosenberg's view of reduction and explanation and reject these additional assumptions that link ontological reductionism so tightly to explanatory reductionism. If my criticisms are convincing they will show that ontological reduction(ism) does *not* give rise to explanatory reduction(ism), but rather that ontological and explanatory reduction(ism) often fall apart.

To put it more generally, even if we concede that "we live in one world", that no vital forces or immaterial entities exist, and that any particular biological entity is constituted by nothing but physicochemical entities (token physicalism), we are still not committed to the reductionist claim that these physicochemical entities represent a "*privileged level of description*" (Mitchell 2003, 181) and that, for instance, all explanations must be formulated in physicochemical terms. There is no such "strong mapping relation" (Mitchell 2003, 181) between scientific representations of the world and the features of the world itself. More importantly, ontological reductionism in the shape of token physicalism does not imply the *priority* of any level of representation. Put another way, the fact that a concrete organism is composed of organs and tissues, which are constituted by cells that are in turn nothing but molecules that are organized and interact with each other in a certain way does *not* imply that all theories about the organism or all explanations of all traits of the organism must be given in terms of molecules. Nor does it imply that the organism's behavior must be studied exclusively on the level of its molecular constituents. In short, ontological reduction does *not* enforce epistemic reduction.

2.2.2 Alternative Relations

This is not to say that there exist *no* connections between ontological and epistemic reduction at all. The opposite is true. But what connections these are depends on the kind of epistemic reduction one is concerned with.

In the case of *theory reduction* à la Nagel (1961) or Schaffner (1993) the implementation of theory reduction requires the development of bridge principles (also known as "bridge laws" and recently called "connectability assumptions" by Schaffner 2006, 386), which connect the different terms of the reduced and the reducing theory. And since bridge principles are typically conceived as expressing identities between types or properties, the accomplishment of theory reduction calls for the existence of these identity relations between these types or properties. For example, the reduction of Mendel's inheritance theories to certain molecular biological theories requires, for instance, the development of a bridge principle that identifies the type dominance with its molecular correlate. Hence, it requires that, in the world, there exists such a relation of identity (i.e. ontological reduction) between the property 'being dominant' of alleles and a certain molecular property.[27] The challenges one encounters when one tries to formulate a bridge principle of this kind led some philosophers to step back from the assumption that bridge principles are identity statements (e.g. Dizadji-Bahmani, et al. 2010, 404f) or to abandon the Nagel-Schaffner approach to reduction altogether (see Section 4). However, if one sticks to the Nagel-Schaffner model, theory reduction demands ontological reduction in the form of type identity.[28]

In the case of *explanatory reduction* there exist also several connections to ontological reduction, even if such a strong enforcement of explanatory reduction(ism) by ontological reduction(ism), as Brandon and Rosenberg envision, must be rejected. To begin with, reductive explanation does not demand ontological reduction in the form of type identity since it can target individual instances instead of types (e.g. Chalmers 1996, 46; Kim 2008, 94-96). That way it avoids the problem of multiple realization.[29] However, even if the

---

[27] My vague formulation 'a certain molecular property' expresses the difficulty or impossibility to find such a molecular equivalent of the type dominance.

[28] More precisely, this claim holds only with respect to Schaffner's account. Nagel himself characterized the nature of bridge statements as "factual or material" (1961, 354), which leaves open whether they are identity statements or not.

[29] At this point one may object that explanations in biology only rarely target tokens. Instead, what is typically being explained is, for example, why coleoptiles (i.e. the protective sheaths covering the emerging shoots in monocotyledons such as oats and grasses) grow in the direction of light (type), not why an individual coleoptile grows in the direction of light (token). However, in many cases the problem of multiple realization can also be avoided by targeting narrower types (which is frequently the case in reductive explanations) or by regarding reduction not as a relation of identity, but of constitution or localization.

reductive explanation of a particular biological phenomenon does not require type identity, it nevertheless requires token reduction (understood either as relation of identity or of constitution/localization). For instance, if an individual ATP synthase would not be composed of two regions (F_1 and F_0), which in turn are constituted of different subunits, each of which possesses a specific amino acid sequence and interacts with the others to build a certain spatial structure, the corresponding reductive explanation that explains the behavior of the ATP synthase (i.e. producing ATP) by reference to the amino acid sequences and spatial structures of its regions and subunits would simply be false. In more general terms, a reductive explanation of a concrete behavior of a system, which refers to certain parts of the system and to certain interactions between these parts, would be inadequate if the system were not in fact composed of these parts and if these parts were not in fact interacting with each other in the described way. In sum, reductive explanation requires ontological reduction in the form of token reduction, but not in the form of type identity.[30]

Besides these general connections my analysis of reductive explanation in biology will reveal additional, more specific connections between explanatory reduction and ontological issues (not necessarily ontological reduction). In particular, I think that if one wants to determine what makes an explanation in biology reductive one must primarily consider ontological issues. This is why I call my account of explanatory reduction an "*ontic*" account. It is not the case that I analyze ontological reduction. Rather, my analysis is about a specific kind of epistemic reduction, namely explanatory reduction (i.e. reductive explanation). But my account of epistemic reduction is "ontic" (analogous to the ontic account of explanation; Salmon 1989, 1993, 1998) since, according to my view, what determines whether an explanation is reductive or not is whether it represents certain relations (e.g. 'x is located on a lower level than y' or 'x is internal to y') that *exist in the world* (for further details see Chapter V, Section 6). It is important to notice that these relations in the world that determine the reductive character of explanations are not the same as the relation of ontological reduction, but related to it.[31]

Furthermore, *methodological reduction* involves various ontological issues, too. In Chapter V I identify two respects in which the utilization of reductive methods (like decomposition and separating a system from its environment) can fail and does not yield reductive explanations. First, a system can exhibit complex forms of interactions with its environment that make it difficult or even impossible to separate the system from its environment. Second, several biological systems are far from being "aggregative systems"

---

[30] In his paper on "Reduction and Reductive Explanation" (2008), Kim goes a step further and argues that reduction (i.e. property identity) even *precludes* reductive explanation: "If an M-phenomenon is identical with a P-phenomenon, there seems to be no specifically M-phenomenon that needs to be, or can be, reductively explained." (2008, 94) But this conclusion is based on an *eliminative* view of ontological reduction, which need not be adopted.

[31] For instance, if we consider token reduction understood as a relation of constitution we might spell out this relation as part-whole relation, which in turn can be specified by the two above-mentioned relations 'being located on a lower level than' or 'being internal to'.

(Wimsatt 1986). Rather, they are highly organized and "functionally integrated" (Bechtel/ Richardson 2010, 149). If the complexity of the organization of a system exceeds a certain limit it is possible that this prevents the decomposition of the system into parts. The system is said to be non-decomposable (see Chapter V, Section 1.2.4). Since non-aggregativity, non-decomposability, and non-separability are all features of biological systems *in the world*, the failure of reductive methods can be traced back to ontological issues.[32]

2.3 Why Epistemic Issues Matter Most

To come straight to the point, I think the issue of ontological reduction(ism) is very interesting and would merit thorough investigation by philosophers of biology. However, since its origin in the 1960s the debate about reduction in philosophy of biology has focused almost solely on epistemic issues. And it is important to emphasize that this is not due to the ignorance or lack of interest of philosophers of biology in ontological issues. Rather, (besides historical grounds)[33] philosophers of biology have a *persuasive reason* for this persistent focus on epistemic issues.

The reason they have is quite simple: epistemic reduction, not ontological reduction, is what is of *central concern* to biologists. Ontological reduction at most plays an indirect, secondary role, for instance, in the sense that a certain case of epistemic reduction presupposes the existence of a specific type of ontological reduction (like Nagelian theory reduction presupposes the existence of type identity relations in the world). However, ontological reduction is not subject to consideration on its own. When biologists discuss the value and limitations of reductions as well as the correctness of reductionism in their research field, they focus exclusively on cases of epistemic reduction or, more precisely, on cases of *methodological and explanatory reduction*. The topic of ontological reduction is not of interest to them – apart from vague lip service that they, of course, reject any form of vitalism and except from statements such as "Molecular biologists… do not hold the naive view that complex structures and processes are just sums of their parts" (Fincham 2000, 343), which are not further specified. Biologists simply accept that, for instance, an individual chloroplast is composed of an outer and an inner membrane, the stroma, stacks of thylakoids, molecules of small circular DNA, ribosomes, and other components, which are spatially organized and interact with each other in a certain way. They neither question what kind of ontological relation exists between a particular chloroplast and its parts, nor whether the type chloroplast can be said to be identical to the type 'assemblage of organized and

---

[32] But notice: It is not the failure of ontological token reduction that gives rise to the failure of applying reductive methods and developing reductive explanations.

[33] In the early stages of the reductionism debate in philosophy of biology the ignorance of ontological issues is probably also an inheritance of logical positivism's repudiation of ontological theses in philosophy of science.

interacting chloroplast membranes, thylakoids, ribosomes, etc' (for further evidence see Chapter V).[34]

The questions biologists are rather occupied with are methodological and explanatory issues, that is, how to carry out fruitful research and how to develop adequate explanations for biological phenomena. In the context of reduction these issues amount to questions like "Is the application of reductive methods a permissible and profitable research strategy?" and "Can a reductive explanation of a system's behavior be considered as adequate and if not, then why does it fail to be adequate?" In order to illustrate this point consider an example from actual biological research, namely the phenomenon of virulence of microorganisms like bacteria and fungi (i.e. their degree of pathogenicity). Some microbiologists who study this phenomenon also reflect upon the issue of reductionism (e.g. McClelland, et al. 2005; Casadevall, et al. 2011). In their papers they primarily address the question whether it is appropriate or not to pursue a reductionist research strategy that identifies a set of microbial characteristics associated with virulence independently of each other and independently of the conditions that are present in a susceptible host (McClelland, et al. 2005, 287). A closely related question is whether the resulting reductive explanation of the virulence of a certain microbe adequately accounts for the observed variation of virulence in relation to other factors and to host dependence (McClelland, et al. 2005, 287; Casadevall, et al. 2011, 1). Hence, these microbiologists do *not* wonder whether the microbe type is identical to a set of different macromolecules interacting with each other. Rather, their discussions center on methodological and explanatory reductionism: Is the application of reductive methods (i.e. the simplification of interdependencies of different microbial features and of microbe-host interactions) the most fruitful strategy to study the virulence of microbes? And are reductive explanations the most adequate modes of explaining the virulence of a certain kind of microbe?

In sum, philosophers of biology focus on questions about epistemic reduction(ism) since biologists themselves are "epistemically profligate", whereas they are "metaphysically frugal" (Brigandt/Love 2008, Section 5). Thus, the *second lesson* one can learn from the previous debate is that *it is epistemology that matters most* – at least if one seeks to understand what reduction in actual biological practice is (see Chapter I).[35] My project of analyzing reductive explanations in biology lines up with this tradition since it is focused on cases of epistemic reduction, too. However, the fact that epistemic questions are the issues that matter most does not imply that questions about ontological reduction are completely

---

[34] Some biologists even explicitly confess that they shy away from addressing ontological questions like these. For instance, Ferric Fang and Arturo Casadevall admit that they were "feeling increasingly uncomfortable" as they "tiptoe[d] gingerly through metaphysics" (2011, 1401) in their paper.

[35] By the way, just as philosophy of biology can learn much about ontological reduction from philosophy of mind (see Section 2.1), it seems to me that philosophy of mind can learn much about epistemic reduction (e.g. what reductive explanations are) from philosophy of biology. Some philosophers of mind have already started to recognize that (e.g. Gillett 2007; Godfrey-Smith 2008).

irrelevant. Accordingly, in this section I hope to have shed some light on what the type of ontological reduction(ism) is, which underlies the debate about epistemic reduction(ism) (Section 2.1), and what kind of connections exist between, on the one hand, epistemic reduction and, on the other hand, ontological reduction, respectively, ontological issues in general (Section 2.2). From here we can now fully attend to the topic of epistemic reduction and, as a first step, distinguish different kinds of epistemic reduction.

3 Third Lesson: Tell Apart Different Types of Reduction

The third lesson one can learn from the previous debate is that before one can give an answer to questions like “Can a certain biological phenomenon be reduced to physical entities?” one needs to know what exactly this question amounts to. In the preceding section we have already become acquainted with two readings of this question. On the one hand, it could refer to *ontological* reduction and mean something like “Can a certain biological phenomenon in the world be said to be identical to or constituted by physical entities?” On the other hand, it could be an *epistemic* question, which does not concern a relation between entities in the world, but a relation between our knowledge or our representations of biological and of physical phenomena. In this case the question could be specified as follows: “Can the knowledge about or the representations of a certain biological phenomenon be reduced to the knowledge about or representation of physical entities?” However, the epistemic question is rarely asked in this general (and quite vague) fashion. Rather, it is either formulated as a question about the relation between *theories* (“Can a biological theory about a certain phenomenon be reduced to physical theories?”), or as a question about *explanations* (“Can a certain biological phenomenon be adequately explained exclusively by reference to physical entities?”), or as a *methodological* question (“Can a biological phenomenon be fruitfully investigated exclusively on the physical level?”). These three questions (i.e. about theories, explanations, and methods) are *subtypes* of the epistemic question. In sum, the above question “Can a certain biological phenomenon be reduced to physical entities?” can be specified in *four* different ways. And depending on how it is specified, completely different answers may be given to it. Thus, the third lesson to learn is that before you engage in the dispute about reduction(ism) you need to specify about which *kind of reduction* you are talking – about ontological reduction, theory reduction, explanatory reduction, or methodological reduction (the latter three of which are subtypes of epistemic reduction).

Philosophers of biology have recognized early on that the issue of reduction comprises different kinds of questions. Francisco J. Ayala (1974) was the first who explicitly introduced the distinction of *different types of reductionism* into the debate. In his introduction to a volume on “Studies in the Philosophy of Biology – Reduction and Related Problems” (1974), which he edited together with Theodosius Dobzhansky, he distinguished three kinds of reductionism: ontological reductionism, methodological reductionism (which he conceived as including explanatory issues), and epistemic reductionism (which he identified with theory reductionism, as was common practice in the early stages of the reductionism debate). Since then, his classification has been taken up and revised (or replaced by a similar one) by many philosophers. For example, some authors add new subcategories (e.g. theory and explanatory reductionism as subtypes of epistemic reductionism; Sarkar 1992; Brigandt/Love 2008, Section 1) and thereby sometimes ignore other categories (e.g. methodological reductionism; Mayr 1988; Sarkar 1992). Others introduce new names (e.g. 'constitutive reductionism' instead of 'ontological reductionism'; Mayr 1988, 2004; Sarkar 1992), and still others employ a strongly modified classification (e.g. ontological and explanatory reductionism are subsumed under the same category; Brandon 1996; see Section 2.2). The distinction of different kinds of reductionism has even made its way into biology itself (e.g. Fang and Casadevall 2011, 1401; Mazzocchi 2008, 11), although the number of biologists who have picked it up remains quite low.

In what follows I refine the distinction between ontological and epistemic reduction that I presented in Section 2 by identifying *three subtypes of epistemic reduction*, namely theory, methodological, and explanatory reduction. What distinguishes these three kinds of epistemic reduction from another is that each of them treats reduction as being a relation between or a feature of *different epistemic units* (i.e. theories, methods, or explanations). My main goal in this section is to further clarify what these three kinds of epistemic reduction are. I will do that, on the one hand, by drawing on results from the previous debate about different types of epistemic reductionism and, on the other hand, by breaking new ground. For example, I specify methodological reduction by focusing on reductive methods rather than on the normative claim of a methodological reductionist (Section 3.2). And I explicate explanatory reduction by examining the reductivity of individual explanations rather than conceiving reduction as a relation between higher-level and lower-level explanations (Section 3.3; see also Chapter III and IV). But first, let us start with examining the most "conservative" kind of epistemic reduction, which has dominated the debate for a long time, namely theory reduction.

3.1 Theory Reduction

When it comes to the issue of epistemic reduction, Ernest Nagel's *The Structure of Science* (1961) constitutes the *locus classicus*. Nobody who writes about reduction in biology fails to mention Nagel's classical model of theory reduction – even if this serves only to distance oneself from Nagel's approach. This is due to the historical fact that the application of Nagel's model of theory reduction to biology (in particular, to the relation between Mendelian genetics and molecular biology; see Hull 1974; Schaffner 1974a; Kitcher 1984; Rosenberg 1985; Waters 1990) marks the *point of origin* of the reductionism debate in philosophy of biology. What is more, it is also one of the few debates that indicated the emergence of philosophy of biology as a distinct discipline (Griffiths 2007). Reasons enough to have a closer look at Nagel's account and at the problems it encountered.

In his *The Structure of Science* (1961) Nagel develops a *formal model of theory reduction*. In the spirit of logical empiricism, Nagel characterizes reduction as a deductive relation that holds between scientific theories, which he takes to be sets of law statements. In line with the deductive-nomological (D-N) model of explanation (Hempel/Oppenheim 1948), Nagel conceives reduction as a special case of explanation. For reduction to occur two conditions must be satisfied: first, the reduced theory has to be derived from the reducing theory ("*condition of derivability*"; 1961, 354). Second, this presupposes that the reduced and the reducing theory either contain the same terms (in case of homogenous reduction) or that the former can be connected to the latter (in cases of heterogeneous reduction) via bridge principles or "connectability assumptions" (Schaffner 2006), which are in the literature also known as bridge laws ("*condition of connectability*"; 1961, 354).[36] At this point it should be acknowledged that Nagel contributed much more to the debate about reduction than this. For instance, he also proposed several non-formal conditions for distinguishing trivial from non-trivial cases of reduction (1961, 358-366), discussed the issues of emergence (1961, 366-380) and "mechanistic explanation" in biology (1961, 398-446), and identified different reasons why the whole can be more than the sum of its parts (1961, 380-397; see also 1952). Nonetheless, the subsequent debate about Nagel's account focused on the *formal conditions* he identifies in his chapter on theory reduction (1961, 336-358).

Although Nagel developed his formal model solely on the basis of examples from physics (in particular, the reduction of thermodynamics to statistical mechanics), the early philosophers of biology regarded it as an adequate understanding of epistemic reduction in the sciences in general and, thus, tried to apply it to biology. In the course of this application it quickly became clear that Nagel's account not only had to face many general problems,[37] but that biology provides *special obstacles* as well. Briefly speaking, the objection was that neither the bridge principles that are needed to connect the terms of biological and physical theories nor the law statements that constitute the relata of theory reduction are available in biology. Let us consider these two objections in more detail.

First, because evolution by natural selection is blind to structural differences with similar functions, most existing biological types of entities are *multiply realized* on the

---

[36] With respect to the nature of these bridge principles Nagel discusses three options (1961, 354-358): either they express meaning equivalence, or they are mere conventions, or they are factual statements. Nagel chooses the last option. However, as Dizadji-Bahmani, et al. (2010, 403-407) have pointed out, this does not commit Nagel to the claim that bridge principles express identity statements.

[37] For instance, Frederick Suppe (1977), Waters (1990) and others criticized the reliance of Nagel's account on a syntactic view of theories. Paul Feyerabend (1962) attacked Nagel's model by claiming the incommensurability of the meaning of the theoretical terms of the reduced and reducing theory. Finally, Schaffner (1967, 1969) pointed out that in most cases of theory reduction the reduced theory first needs to be corrected before it can be derived from the reducing theory.

physical level.[38] For example, the wings of different species of birds (let alone those of mammals and insects) vary strongly with respect to their structure and material composition although (almost) all of them share the same function, that is, they enable their bearers to fly. The multiple realization of biological types makes it very difficult or even impossible to establish those connections between the terms of biological (e.g. classical genetics) and physical or molecular theories (e.g. molecular biology) that are needed for theory reduction in the Nagelian sense. Second, another obstacle for a neat application of Nagel's model to biology was his assumption that theories are sets of *law statements*. The generalizations that can be found in biology (e.g. Mendel's laws of segregation and independent assortment) seem to be far away from describing laws of nature in the classical, strict sense. They typically have exceptions, are restricted in scope, and arguably are historically contingent (Beatty 1995). This led many philosophers of biology to the conclusion: no laws in biology, no theories in biology, hence, no cases of reduction in biology.

The result of these problems was the formulation of the "antireductionist consensus" (Waters 1990, 125). About 20 years after the reductionism debate in the philosophy of biology had emerged it seemed as if everybody had become an antireductionist.[39] Even philosophers with strong reductionist intuitions like Rosenberg gave up the hope that biology could be reduced to physics.[40]

It is important to note that during these 20 years and up to the 1990s the majority of philosophers took the obstacles to applying Nagel's model to biology to reveal the non-existence of reduction in this field and to support the incorrectness of reductionism in biology. Most of them did not choose the alternative option to question that Nagel's account is, in principle, the adequate way of thinking about epistemic reduction.[41] Despite this general agreement, it was common practice to disagree *about the details* of the Nagelian model of theory reduction and to call for revisions. Several philosophers of biology tried to overcome the problems of Nagel's account by developing it further. Most notably, Kenneth Schaffner (1969, 1974a, 1974b, 1993) developed his "General Reduction-Replacement (GRR) Model" (1993, 429). With his changes Schaffner addresses a key drawback of Nagel's model, namely that the reduced theory may contain empirically false statements and, in this case, cannot be deduced from the reducing theory (a critique originally raised by Feyerabend 1962). In order to cope with this problem Schaffner allows that the reduced theory is *corrected* before being reduced (whereby the original reduced theory and its corrected version need to be "strongly analogous" 1993, 429).

---

[38] For a detailed elaboration of this point see, for instance, Rosenberg 2001.

[39] Notable exceptions are Ruse (1976) and Schaffner (1969, 1976).

[40] In his 1994 book Rosenberg argues that the impossibility of reductionism in biology inevitably leads to an instrumentalist interpretation of biological theorizing and to the abandonment of the unity of science above the level of physics. However, in the 2000s Rosenberg gave up this antireductionist position and became one of the few contemporary defenders of reductionism in biology again.

[41] Among the few exceptions were Wimsatt (1976a) and Hull (1974).

However, despite these modifications, at that time hardly anybody questioned that Nagel's model is, in general, the adequate way of thinking about epistemic reduction. That is, most philosophers accepted the following two *core assumptions* of Nagel's account:

(1) The adequate units of the relation of reduction are *theories* (whether they are conceived as sets of law statements or not, whether the theories need to be corrected before being reduced or not, and whether one adopts a syntactic view of theories or not).[42]

(2) The relation of reduction is a relation of *logical derivation* (whether this means exact derivability or something weaker and whether the bridge laws that are necessary for derivation are conceived as identity statements or not).[43]

The widespread acceptance of this general way of thinking about reduction in terms of theories and relations of logical derivation prevailed in the debate for a surprisingly long time.[44] The most instructive example is Rosenberg, who nowadays explicitly argues for the need to abandon the Nagelian understanding of reduction (2006, 40) but, in the 1980s and 1990s, claimed that it "sounds suspicious to change the standards of reduction" (1985, 110) and conceived the alternative option of abandoning reductionism altogether as the "more reasonable" (1994, 22) option.

But the situation has changed. Since the 1990s more and more philosophers reject even the core assumptions of Nagel's approach and develop alternative ways of thinking about epistemic reduction in biology (e.g. Sarkar 1992, 1998, 2005; Wimsatt 1976a, 2007; Rosenberg 2006; Bechtel 2006, 2008). My account of explanatory reduction in biology lines up with this tradition since it proposes a new way to understand what epistemic reduction in biology is. As such, it also tries to establish clarity in the "polyphonic disunity" (Wimsatt/Sarkar 2006, 697; Wimsatt 2006a, 447) that has been left after the abandonment of Nagel's "unitary" account. With my project of developing an account of explanatory reduction I choose a different path than the one several other contemporary opponents of the Nagelian approach run. Instead of elaborating an alternative view of epistemic reduction they argue for the *abandonment* of the focus on reduction altogether and propose a different relation between fields, for instance, "integration" (Craver 2005, 2007a; Mitchell 2003, 2009; Mitchell/Dietrich 2006; Darden 2005; Darden/Maull 1977). Contrary to them, I think it is important to adhere to the concept of epistemic reduction because it is an

---

[42] Although some philosophers questioned the syntactic view of theories and called for a less formal alternative, up to the late 1990s almost nobody questioned the general thesis that *theories* are the adequate units of reduction. For instance, in his influential paper from 1990, Waters objected to Nagel's model of theory reduction but merely demanded the "reformulation of theoretical reduction" (1990, 136). Nowadays he explicitly criticizes the concepts of "theoretical reduction" and "layer-cake antireduction" (Waters 2008, 239) and the exclusive focus on *theoretical* developments in biology they imply.

[43] At this point I should stress that there in fact were a few philosophers of biology (most notably, Hull 1976 and Wimsatt 1976a) who early objected to this second thesis, that is, Nagel's and Schaffner's presupposition that a model of theory reduction should focus on formal issues and reconstruct reduction as a relation of logical derivation.

[44] This is especially true for discussions that are not centered on but rather pick up the issue of reduction.

important conceptual tool for capturing significant aspects of biological research practice. What we need is a new account of what epistemic reduction in biology is, not the removal of the concept of reduction from philosophy of biology altogether.

In Section 4 I will pick up the issue of theory reduction again and fill in the argumentative step that is missing so far. I will argue why it is time to move beyond Nagelian reduction and to give up even the two core assumptions of Nagel's approach presented above. This should convince even the last contemporary proponents of Nagelian reduction (e.g. Schaffner 2006; Bickle 1998, 2003, 2006; Krohs 2004; Klein 2009; Dizadji-Bahmani, et al. 2010; etc.) that this model *fails* to capture what epistemic reduction in biology really is. On this basis we can then, in Chapters III to VI, begin to search for a new account of epistemic reduction in biology. But first, let us continue to specify the different kinds of epistemic reduction one needs to keep apart.

3.2 Methodological Reduction

The notion of methodological reduction is not common to the debate – contrary to the concept of methodological reductionism. However, I will argue in the following section (3.2.1) that disputes about methodological reductionism require an understanding of what reductive research strategies or methods in biology are. In other words, discussions about methodological reductionism demand a concept of methodological reduction. In Section 3.2.2 and 3.2.3 I provide first steps in developing such a concept. In Chapter V I pick up on this task again, identify and analyze the major reductive methods that are applied in contemporary biological practice.[45]

3.2.1 Wannabe Reductionism

In the recent debate methodological reductionism has been regarded as a thesis about the *most fruitful* way of practicing biological research. Methodological reductionists claim that we *should* "always seek explanations by investigating the underlying processes at lower levels of complexity, and ultimately at the level of atoms and molecules" (Ayala 1974, viii). Thus, proponents of methodological reductionism endorse a *normative* claim or, in other words, put forward a "procedural recommendation" (Brigandt/Love 2008, Section 1). Their argument is that research should be pursued in a reductionist fashion because this is the most successful way to pursue it.

Typically, the phrase 'practicing research in a reductionist fashion' is spelled out with reference to *reductive explanations*: what makes research strategies reductive is that they *aim at* developing reductive explanations. For instance, Rosenberg characterizes methodological reductionism as the "methodological moral that biologists should seek…

---

[45] I treat methodological reduction as a subtype of epistemic reduction since I think that investigative strategies, heuristics, or methods are parts of the knowledge of a scientific discipline (more precisely, they are part of what is called "knowing-how").

macromolecular explanations" (2006, 26). Likewise, Schaffner states that a "reductionist research program… will attempt to explain biological processes in terms of physical and chemical sequences of events" (1974a, 127). This does not mean that these authors make the mistake to overlook the difference between methodological and explanatory reduction (a difference that is, for instance, emphasized by Sarkar 1992, 169; 2005, 106). Rather, most of them recognize this difference.[46] But the only way they characterize what it amounts to conduct biological research in a reductive manner is by reference to the unspecific phrase 'seeking reductive explanations'.

Wimsatt takes a radical stance towards this widespread vague notion of methodological reductionism. He accuses methodological reductionists of practicing "*wannabe reductionism*" (2006a, 445; my emphasis). That is, they claim that one should pursue reductionism, but never propose *how*.

> What is called 'methodological reductionism' in the philosophical literature could better be named 'wannabe reductionism'. It appears to be the view that we don't know whether reductionism is correct, but let's pursue our research as if it were. Fine! But then we are never given any hints as to how we should act in the laboratory, or what strategies we should follow in building our models. And these writers appear to have no interest in finding out. (Wimsatt 2006a, footnote 1)

I agree with Wimsatt on the point that the widely accepted notion of methodological reductionism does not sufficiently specify what reductive research strategies or methods in fact are (or should be) and, thus, leaves a "bothersome lacunae" (Wimsatt/Sarkar 2006, 697). Since we want philosophy to be connected to what scientists actually are doing and to be potentially useful for practicing scientists (see Chapter I, Section 6) this gap needs to be closed. Hence, in the subsequent sections I first spell out the notion of a reductive method (Section 3.2.2). Then, I specify the relation between reductive methods and reductive explanations (Section 3.2.3). And finally, I address the putative objection that reductive methods, not reductive explanations, are the appropriate target of an analysis of epistemic reduction (Section 3.2.4).

3.2.2 What Are Reductive Methods?

My aim in this section is to provide some clarification about what it means to pursue biological research in a reductive fashion. That is, I take the first steps towards a specification of the notion of a reductive method. In order to do so I pursue two goals. First, I explicate what the term '*method*' refers to by specifying its scope. Second, I turn to the *reductive character* that a method can possess. I spell out what determines the reductivity of a method by analyzing different reductive methods that can be found in biological practice.

---

[46] Notable exceptions are, for example, Ayala who fails to identify explanatory reductionism as a distinct category and intermingles methodological and explanatory issues under the one label 'methodological reductionism' (1974, viii-x).

Reductive Methods as Strategies or Heuristics

Let us start with the first step. The term 'method' can refer to a variety of different elements, ranging from very specific (laboratory) techniques over investigative strategies or "heuristics"[47] (Wimsatt 2006a, 463; 2007, 76; Bechtel/Richardson 2010, xxx) to general epistemic stances. It seems to me that the best way of specifying the notion of a method – at least in the context of reduction – is to identify them with such middle range elements like *investigative strategies* or *heuristics*. But which reasons can be adduced for this decision?

It seems to me that there are (at least) two alternatives: one could regard reductive methods as specific techniques or as global stances. Consider first the alternative that reductive methods can be characterized as small-range methodological elements like laboratory *techniques* (e.g. PCR, Southern blot, screening, etc). The disadvantage of this choice is that most techniques are mere instruments for collecting empirical data and as such too discipline-specific and too diverse to be a proper subject of a philosophical analysis of the concept of a reductive method. The second alternative, that reductive methods are global elements, such as entire reductionist *approaches, stances or attitudes*, is a serious alternative because this is what some biologists mean by 'reductionism'. For example, in cancer research the putatively misguided "reductionist stance" (Soto/Sonnenschein 2009, 3) is opposed to the integrative, "organicist approach" (Soto/Sonnenschein 2009, 6). Since the former directs attention to the discovery of the "faulty" genes that cause cancer as well as to intracellular processes it precludes the investigation of tissue level causes. According to the proponents of the organicist approach, precisely these neglected tissue-level causes are most important for explaining cancer. However, characterizing reductive methods as global reductionist stances has important drawbacks. It suggests that the choice between applying reductive methods or not is an either-or choice. The reason is that taking up a global reductionist stance is incompatible with adopting a non-reductionist approach at the same time. But exactly this, namely investigating a phenomenon by *applying reductive and non-reductive methods simultaneously* (or in succession), is what several biologists seem to regard as "good research". They do not argue against the utilization of reductive methods in their research area *per se* and demand their complete elimination.[48] Rather, they want to "broaden the picture" (Powell 2004, 299) and emphasize the need to supplement reductive methods with more holistic ones (or with an "integrative agenda" Gallagher/Appenzeller 1999, 79). In their view "knowing the parts, and even knowing the function of the parts, is not enough." In addition, we need to "simultaneously stud[y] the complex interaction of many levels of biological information" in order to "understand how they [i.e. the parts] work together." (Keller 2005, 5)

---

[47] For a general characterization of the concept of a heuristic see especially Wimsatt 2007, 76-84, Appendix A and Bechtel/Richardson 2010, xx-xxviii. Various common reductive heuristics are listed in Wimsatt 2007, Appendix B and in 2006a, 467-472.

[48] Even self declared anti-reductionists as for instance Ernst Mayr acknowledge the value of the reductive method. He calls it "analysis" and argues that it is "one of the most successful methods in science" (1988, 475).

Hence, reductive methods are best understood as being mid-range reductive heuristics or investigative strategies (rather than as being specific laboratory techniques or as global reductionist stances), which can be combined with the application of non-reductive research strategies. But what exactly are these reductive heuristics or strategies? This question is examined in the subsequent sections by considering paradigmatic examples of reductive methods.

Decomposition and Other Reductive Methods

Let us turn to the second step, that is, to the question of what distinguishes reductive methods from non-reductive ones. The reductive method *par excellence* is the "dissection of biological systems into their constituent parts" (van Regenmortel 2004, 1016), which is also known as "*decomposition*" (Bechtel/Richardson 2010, 23; my emphasis), "analysis" (Ayala 1974, vii; Mayr 1988, 475), or "downward looking" (Lidicker 1988, 278; Byerly 2003, 337; Bechtel 2009, 543). Consider the example of protein synthesis in eukaryotes. To investigate this phenomenon in a reductive manner means to decompose the corresponding system, namely the cell, into certain parts (like ribosomes, the genome, amino acids, m- and t-RNAs, splicing enzymes, etc.) and to study the behavior of these parts. Although other reductive methods can be identified as well, decomposition occupies a special status since it is by far the most prominent and most discussed reductive method in the biological sciences. The method of decomposition can be further qualified in two respects:

First, it is important to note that an individual biological system as well as different biological systems can be decomposed differently, that is, into different kinds of parts (see Chapter V, Section 1.2.1) and into parts that are located on different organizational levels (see Chapter V, Section 2). The level of decomposition, that is, the level on which an explanation "bottom[s] out" (Machamer, et al. 2000, 13) depends on which kind of parts and interactions/activities are taken to be unproblematic and relevant to the explanatory interests and purposes of a given scientific field (see Chapter V, Section 1.2.3). Even though the decomposition of many biological systems identifies molecules and their interactions as genuine parts, the decomposition need not bottom out at this "fundamental level" (Chapter V, Section 2.3.1). For instance, the dissection of a kidney into tissues and cells of different kinds (e.g. into nephrons, fibrous tissue, the renal capsule, etc.) is as much an example of the usage of a reductive method as the dissection of a transfer RNA (tRNA) into a certain sequence of typically 73 to 93 nucleotides (which are macro*molecules*). This fact is disregarded by philosophers of biology who specify the notion of a reductive method in such a way that only the "search for *molecular* completions, corrections, or foundations of… functional explanations" (Rosenberg 2006, 26; my emphasis) counts as an application of a reductive method.

Second, two subtypes of decomposition can be distinguished, namely *structural (or spatial) decomposition* and *functional (or temporal) decomposition*. This difference captures the fact that biological systems exhibiting a particular behavior (e.g. the transformation of light energy into chemical energy in chloroplasts) can be decomposed into two kinds of constituents. On the one hand, they can be spatially decomposed into certain objects (e.g. thylakoid membranes with photosystems, electrons, ATP-synthases, proton-gradients, cytochromes, etc.). On the other hand, they can be temporally decomposed into certain operations, interactions, activities, or however you refer to what the spatial parts of a system are doing, that is, how their properties are changing over time (e.g. the transport of electrons, the creation of a proton gradient across the membrane, the absorption of a photon by a photosystem, etc.). Philosophical discussions about decomposition have focused on spatial issues and on parts as objects for a long time. However, biological systems do not only consist of static objects with certain unchanging properties (see Chapter V, Section 1.2.2). Rather, these objects interact with each other and change their properties. You may also describe this as the objects being engaged in specific “activities” (Machamer, et al. 2000) or performing certain “operations” (Bechtel/Richardson 2010; Bechtel 2006, 2008). It is important to note that the distinction between structural and functional decomposition does not commit me to the much stronger ontological thesis that spatial and temporal parts represent two *distinct ontological kinds*, as for example Machamer, et al. (2000) and, in particular, Machamer (2004) have argued. By contrast, I endorse the more moderate claim that the difference between the spatial and temporal parts of a system is merely *descriptively adequate*. Otherwise put, I hold that this difference captures an important distinction biologists actually make when they conceptualize, represent, and reason about the constituents of biological systems (and mechanisms). Whether this difference points to an ontological dualism of entities and activities or whether activities are nothing more than interactions (i.e. changes of properties) is a question I remain agnostic about.

The difference between structural and functional decomposition might give rise to a second reductive strategy, which Bechtel and Richardson have called *localization*” (2010, xxxii, 24).[49] According to their view, decomposition involves the partitioning of the behavior of the system to be explained into smaller units, that is, into those component *operations* (or functions) that in sum yield the behavior in question.[50] Afterwards, these operations need to be spatially localized in different parts of the system. In other words, one needs to show that “*something* is performing each of these functions” (2010, 24). This reductive strategy of

---

[49] The huge discrepancy between the standard philosophical framework of theory reduction and the picture of reduction that can be found in the life sciences themselves led Bechtel and Richardson to avoid talking about reduction at all. Accordingly, they also shy away from calling decomposition and localization *reductive* strategies. However, in the preface to the original edition they explicitly state: “the attempts to localize the causes of phenomena in components of complex systems… reflect what many scientists understand by the term *reduction*” (2010, xiii).

[50] In doing so, the “explanatory task… becomes manageable and the system intelligible” (2010, 23) since the behavior of the whole system is presented as a set of subordinated activities or functions performed in the system.

"mapping the operations into which the overall function of the mechanism [or system] is decomposed onto the parts into which the structure is decomposed" (Bechtel 2006, 32) is termed "localization".[51]

So much for the most prominent and widely discussed reductive method, namely decomposition. But what other reductive methods are employed in biological practice? A possible answer can be found in the work of Wimsatt. Most notably in his recent work (2006a, 466-472; 2007, 80-84, Appendix B), Wimsatt presents an entire list of different "common reductionistic heuristics" (2007, 347), which he divides into groups depending on the kinds of activities or contexts that occasion their use (e.g. conceptualization, model-building, experimental design, etc.). Prior to this list, Wimsatt characterizes the procedure of a reductionist in general terms. In his view, reductionist strategies focus our attention on the parts of a system, on their properties, their interrelations, and their interactions.[52] Accordingly, reductionists "ignore or downplay… the importance of the environment [of the studied system]" (Wimsatt 2006a, 466). In particular, this amounts to, for instance, the following reductionistic heuristics:[53]

- "(3) *Interface determinism*:… *black-world perspectivalism* – all that matters about the environment is what comes in across system boundaries" (2006a, 468)
- "(4) *Entificational anchoring*: Assume that all descriptions and processes are to be referred to entities at a given level… assuming that there is a single cause for a phenomenon" (2006a, 368)
- "(5) *Modelling localization*: look for an intra-systemic mechanism to explain a systemic property rather than an inter-systemic one… *Structural* properties are regarded as more important than *functional* ones" (2006a, 468)
- "(6) *Contextual simplification*:… simplifying environment before simplifying system" (2006a, 468)
- "(7) *Generalization*:… focus on generalizing or elaborating the internal structure, at the cost of ignoring generalizations or elaborations of the environmental structure" (2006a, 469)
- "(10) Locality of *testing*: Test a theory only… under laboratory conditions… rather than testing it in natural environments" (2006a, 469)
- "(13) *Articulation-of-Parts (AP) coherence*: Assuming that the results of studies done with parts studied under different… conditions are *context-independent*, and thus still valid when put together to give an explanation of the behavior of the whole." (2006a, 470)
- "(19) *Extra-perspectival blindness or perceptual focus*: Assuming that a system can be exhaustively described and explained from a given perspective… This bias interacts with #4 to give *extra-level blindness*" (2006a, 471)[54]

---

[51] Bechtel and Richardson do not always keep decomposition and localization conceptually apart as neatly as one might wish. For instance, sometimes they seem to include the identification of the spatial parts of a system (or mechanism) in the decomposition task (2010, 26) and sometimes in the localization task (2010, 24).

[52] This requires that the considered system is first decomposed into parts.

[53] Since the list of common reductionistic heuristics that Wimsatt presents in the appendix of his book (2007, 347-352) and in Section 11 of the 2006a paper are identical, in what follows I will only refer to the 2006a paper.

[54] The reductionist heuristics (3) and (4) belong to the first class of "Biases of conceptualization" (2006a, 468); (5), (6), and (7) are grouped under the label "Biases of model-building and theory construction" (2006a, 468); (10) and (13) are classified under "Observation and experimental design" (2006a, 469); and (19) belongs to "Other important biases" (2006a, 471).

Here I have only cited the essential passages from what I take to be the most significant reductionistic strategies Wimsatt presents. When I develop my account of reductive explanation in Chapter V I come back to them again and discuss them in detail one by one. For now, this overview of which kind of reductive methods (besides decomposition) can be identified in biological practice should suffice.

As a last point, let me make you aware of a difficulty in Wimsatt's otherwise impressive work on reductive methods. This difficulty is the main reason why I think not all reductionist strategies Wimsatt lists are convincing and why I have excluded some of them from the above quote. The problem arises because Wimsatt identifies reductive methods with "biases" (2006a, 468), "fallacies" (2006a, 470), and characterizes them as "error[s]" (2006a, 470). This results, on the one hand, in a too negative picture of reductive methods since "[r]eductionistic science is not all bad" (Hull/van Regenmortel 2002, 12). On the other hand, it misleads Wimsatt to classify several problematic research strategies as reductive although it is highly doubtful that they in fact are reductive.[55] Examples of these problematic but not reductive strategies are: "(11) *Abstractive reification*: Observe or model only those things common to all cases; don't record individuating circumstances" (2006a, 469), "(18) Imposition of *incorrect* set of *functional* categories" (2006a, 471), or "(20) *Tool-binding*: Becoming sufficiently bound to a specific… tool" (2006a, 471). However, if one keeps this error in mind Wimsatt's work nevertheless provides a rich source of methods that can properly be called reductive.

3.2.3 The Relation between Reductive Methods and Reductive Explanations

As we have seen in the previous section, the notion of a reductive method can be specified in a substantial way that goes beyond merely characterizing reductive methods as strategies for seeking reductive explanations. Methodological reductionism need not be wannabe reductionism. However, despite the fact that there is a rich notion of reductive method available one may still insist that there is a *close link* between on the one hand reductive methods and on the other hand reductive explanations. Isn't it the case that the study of a phenomenon by applying reductive methods automatically leads to the development of a reductive explanation of this phenomenon? For instance, if a biologist investigates the behavior of a certain biological system by decomposing it into parts and studying the parts in isolation won't the outcome of this process, the explanation of the phenomenon, inevitably be a part-whole explanation and thus reductive, too?

I agree that reductive methods and reductive explanations are two elements of biological practice that are closely connected with each other. This is the reason why we will

---

[55] One might wonder on which basis I assess these strategies as being non-reductive. On the one hand, this judgment is already influenced by the result of my analysis of reductive methods and reductive explanations that I present in Chapter V. On the other hand, I think it is implausible and inadequate to biological practice to draw the difference between reductive and non-reductive methods in such a way that almost all strategies count as reductive. However, Wimsatt runs the risk of doing exactly this.

encounter the issue of reductive methods in Chapter V again, although this chapter is targeted at developing an account of reductive explanations in biology. Anyhow, I think it is important to emphasize that explanations and methods are two *different* epistemic elements of biological practice that need to be kept apart: Explanations can be characterized as the *result* (or outcome) of the *process* of applying certain methods. Thus, they are closely connected, but not the same. More importantly, there are clear cases, in which the application of a reductive method does *not* come along with the development of a reductive explanation.

First, there exist biological fields, so called data-driven sciences (e.g. genomics and parts of systems biology), where the application of methods is not (yet) accompanied by the formulation of hypotheses and the construction of explanations. In other words, methods can be employed *without developing explanations*. At least, methods need not be constrained by explanatory aims but can be open-ended about explanation (e.g. in case of explorative experimentation).

Second, even if the primary aim of using certain methods is to construct an explanation for the investigated phenomenon, the fact that (some of) the applied methods are reductive does *not guarantee* that the *resulting explanation* is *reductive*, too.[56] On the one hand, the application of reductive methods can fail altogether – for instance because the studied system is non-decomposable into parts or non-separable from its environment (in this case the utilization of this reductive method alone probably yields no explanation at all). On the other hand, the use of reductive methods can fail since it results in inadequate reductive explanations – for example if studying the parts in isolation does not shed light on the relational properties the parts exhibit *in situ* or if the environment affects the studied system in a way that cannot be ignored or simplified. However, in the second kind of cases it is also possible that the application of reductive methods only fails to provide an adequate *reductive* explanation of the investigated phenomenon, but that it succeeds to yield an adequate *non-reductive* explanation (or at least that it yields some important steps towards it). The reason is that “we can learn an immense amount even from unsuccessful attempts at reduction” (Ayala 1974, xv; see also Popper 1974; Wimsatt 1976a, 2006a; and Brigandt/Love 2008) – for instance, about the importance of environmental conditions for the behavior of the studied system, about the conditions when aggregativity fails, etc. In cases like these the failure of the reductivity of the explanation does not imply a failure of explanation *per se*. All in all, the reductivity of explanations is *partially independent* from the reductive character of the applied methods.

However, in many cases the exclusive usage of reductive methods *indeed* gives rise to the development of reductive explanations, which often are accused of being inadequate (especially when the behavior of complex systems is concerned). This state of affairs led philosophers of biology to call for the application of a *diversity of methods* – reductive

---

[56] Sarkar (2005, 119f) makes a related point.

together with non-reductive ones (e.g. Mitchell 2009; Brigandt/Love 2008, Section 5; Bechtel/Richardson 2010; Bechtel 2008, 148-157). Likewise, biologists demand to "broaden the picture" (Powell 2004, 299) and to supplement reductive methods with more holistic ones (or with an "integrative agenda" Gallagher/Appenzeller 1999, 79; see also Section 3.2.2). For example, some biologists study the functioning of the heart by decomposing it into certain parts and by investigating the behavior of these parts. But in addition to that, "protein interactions within the context of subcellular, cellular, tissue, organ, and system structures" are also computationally modeled in order to "determine the logic of healthy and diseased states" (Noble 2002, 1678). That way the value of reductive methods is acknowledged, but not overestimated, since the need to supplement them with non-reductive methods is recognized.[57] As Wimsatt puts it: "We need our [reductionist] heuristics, but must use them… with special care" (2006a, 461).

3.2.4 Are Methodological Issues the Ones that Matter Most?

Given the significance of methodological considerations in biological practice in general and in biologists' disputes about reductionism in particular the question arises why we should not join authors like Wimsatt and Waters and focus on reductive methods when analyzing what reduction in biology is. Although Wimsatt originally was also concerned with reductive explanations (1976a) in recent years his attention shifted more and more solely to reductionistic heuristics and their role in science (e.g. 1997, 2006a, 2006b, 2007). According to Wimsatt, this focus on methodological issues arises from the attention philosophers should pay to the details of scientific practice. Rather than analyzing science in terms of flawed idealizations about how scientists reason, philosophers should consider the strategies scientists actually employ (2007, Chapter 1-3). In a similar fashion Waters (2008) criticizes that the present dispute between reductionists and antireductionists is concentrated on questions about the relation between theories and about the adequacy of lower-level explanations. This focus on theories and explanations leads philosophical attention astray and prevents philosophy from understanding "how DNA retooled genetics and transformed biological practice" (2008, 238). The reason is that the development of theories and explanations in genetics is only "peripheral" (2008, 251) to this transformation. In order to reveal what the real developments were and why they were so useful, philosophers need to

---

[57] For those philosophers who want to know whether reductionism or antireductionism is ultimately correct, the question arises whether such a position can still be called reductionist or not. Is a philosopher, who stresses the significance and the past success of reductive methods but simultaneously concedes that in many cases they need to be complemented with non-reductive methods, still a methodological reductionist? The way that the notion of methodological reductionism traditionally has been understood suggests that he is not a reductionist since he does not endorse the unrestricted fruitfulness of applying reductive methods. But it seems to me that such a person cannot be called an antireductionist in the strict sense, too, since he *does* acknowledge the importance of reductive methods. Perhaps such a position is best referred to as neither reductionism, nor antireductionism, but as methodological pluralism.

shift their attention to the way the "investigative strategies" (2008, 257) in genetics have changed.[58]

It should be no surprise anymore that I endorse Wimsatt's and Waters' careful consideration of biological practice (see Chapter I). But despite this general accordance I disagree with their basic assumption that reductive methods are the only proper or the best targets of an analysis of epistemic reduction. I concede that reductive methods are an important element of biological practice and that their philosophical analysis is a valuable project. However, there are at least three reasons why I do not subscribe to Wimsatt's and Waters' emphasis on the importance of reductive methods.

First, Waters' argument does not apply to my analysis of epistemic reduction in biology since my aim is not to understand how biology or a certain biological field like genetics developed or has been transformed over time. In other words, my analysis is concerned with interlevel reduction, not with successional reduction (see also Section 3.4).

Second, it seems to me that from the perspective of biology reductive explanations play an equally significant role in biological research practice as reductive methods. And as such reductive explanations deserve at least as much philosophical attention as reductive methods. From this perspective my choice to develop an account of epistemic reduction by focusing on reductive explanations appears to be as much a *matter of interest* as Wimsatt's and Waters' choice to take into account only reductive methods.

Third, I should emphasize that my focus on reductive explanations is not only due to my personal interests. In addition, I think there exist *good reasons* for a philosopher of biology to turn his attention to the analysis of reductive explanations rather than of reductive methods. Explanations constitute a better object of philosophical study than methods because they are both more constrained and less discipline specific.[59] In general, scientific methods change over time because scientists are quite promiscuous in choosing their methods. In contrast, explanations are more constrained and, thus, more stable (although they, of course, also change). Likewise, the methods that are applied in scientific practice vary more from field to field, that is, they are more discipline specific than explanations. Hence, methods are less comparable than explanations. Thus, reductive explanations constitute a more suitable target of philosophical analysis. However, this does not mean that *no* constrained and transdisciplinary methods can be found in scientific practice, which are adequate objects of philosophical investigation. Moreover, it does not

---

[58] To which of the two layers Waters assigns explanations remains obscure. On the one hand, he clearly argues that focusing on the development of theories *and* of explanations in genetics is wrongheaded. For instance, Waters accuses Wimsatt (1976a) and Sarkar (1998) for their "focus on how genetics explain or try to explain phenomena" (2008, 253). On the other hand, Waters speaks about the shift of attention from "theory" (2008, 239) or "explanatory theories" (2008, 241) to practice. And, more importantly, he includes "explanatory reasoning" (2008, 253) within the investigative practice of science. Perhaps the distinction Waters is up to is between, on the one hand, *explanations* as the result of the utilization of, on the other hand, *explanatory reasoning strategies* (which belong to the class of scientific methods).

[59] Thanks to Alan Love for making me aware of this point.

imply that we do not need analyses of methodological reduction in order to get a comprehensive view of epistemic reduction in biology. The contrary is the case. In this point I agree with Sarkar who states that it is implausible to assume that “all cases of reduction are so similar that they can all be captured by a single model of reduction” (1992, 188).

3.3 Explanatory Reduction

Since the emergence of the reductionism debate in philosophy of biology, epistemic reduction has been *tied closely* to explanation. In accordance with the deductive-nomological (D-N) model of explanation (Hempel/Oppenheim 1948), Nagel characterized theory reduction as a special case of explanation, namely as the explanation of the reduced theory by the reducing theory. Accordingly, when Nagel’s model of reduction was applied to biology (i.e. to the relation between classical genetics and molecular biology) one important question was whether “derivations of the transmission laws from principles of molecular biology and bridge principles… *explain* the laws” (Kitcher 1984, 347; my emphasis) or not. Thus, Nagel’s model of theory reduction is, so to speak, simultaneously a model of explanatory reduction.[60]

As the Nagelian approach was abandoned by more and more philosophers of biology it soon became clear that an alternative understanding of epistemic reduction was needed. But although many authors concurred that reduction is closely related to explanation, no consensus about the proper alternative account of epistemic reduction emerged. Some even describe the actual situation as one in which the disappearance of Nagel’s “unitary account of reduction” has left a “polyphonic disunity” (Wimsatt/Sarkar 2006, 697). In the following chapter I review the most significant proposals concerning the relation between reduction and explanation that have been made so far. I will argue that the proposed accounts of explanatory reduction can be divided into two classes. On the one hand, authors like Rosenberg characterize explanatory reduction as a relation between two explanations, namely between a higher- and a lower-level explanation of the same phenomenon. On the other hand, in the work of Kauffmann, Wimsatt, Sarkar, Hüttemann, and Love, individual reductive explanations like part-whole explanations occupy center stage. In these cases explanatory reduction can be reconstructed as a relation that exists between the phenomenon to be explained and the explanatory relevant factors (for further details see Chapter III; my analysis of explanatory reduction is presented in Chapter V).

---

[60] In order not to confuse these two categories I restrict the notion of an account or model of explanatory reduction to those accounts that abandon the core assumptions of Nagel’s approach (see Section 3.1) and reconstruct reduction primarily as a relation between explanations or between parts of an explanation.

3.4 Successional vs. Interlevel Reduction

This section (3.4) is not fully in line with the other three (3.1-3.3) since it does not address another subtype of epistemic reduction. Rather, it introduces a distinction that is important for the topic of epistemic reduction in general. In fact, this distinction was established in the context of theory reduction, but it seems to me that it can also be conveyed to the context of explanatory reduction. The distinction I allude to is the one between successional reduction and interlevel reduction.[61]

Thomas Nickles was the first who made explicit these "two concepts of intertheoretic reduction" (1973), closely followed by Wimsatt (1976a, 675-679; 1976b, 216-223). Then and now Wimsatt emphasizes the importance of distinguishing these two kinds of reduction since they serve "fundamentally different functions" (2006a, 448) in scientific practice: successional reduction plays a role in theory development and is a domain-preserving relation, whereas interlevel reduction fulfills domain-combining functions.[62] Let us be more specific. *Successional reduction* relates two theories from the same domain (i.e. they typically have the same area of application), namely a historically earlier theory to its successor. A frequently discussed example is the reduction of special theory of relativity to classical mechanics (by taking the limit as velocity approaches zero).[63] If the reducing theory is more mature than its historical predecessor reduction can be an instance of theoretical progress. In the case of successional reduction the successor theory *replaces* or eliminates the predecessor theory only if reduction fails (Wimsatt 1976a, 677; 2006a, 450). In comparison, most philosophers would agree that *interlevel reduction* is never eliminative (see also Section 1). Interlevel reduction is a relation between two theories from different domains (i.e. with different areas of application) that typically are accepted at the same time. In other words, interlevel reduction relates a higher-level theory to a lower-level theory, as for instance the putative reduction of Mendelian genetics to molecular biology.[64]

In line with his general critical attitude against Nagelian models of theory reduction, Wimsatt early argued that in biology interlevel reductions rarely take place between theories, but rather are a matter of explanations (1976a, 675-679).[65] One should notice that the fact that there exist no (or only very few) cases of interlevel theory reduction in biology

---

[61] Several other names for this distinction have been put forward – for example intralevel vs. interlevel reduction or diachronic vs. synchronic reduction.

[62] Nickles (1973) characterizes the different functions as follows: successional reduction (i.e. his "reduction$_2$") serves heuristic or justificatory purposes, whereas interlevel reduction (i.e. "reduction$_1$") provides unification and explanation.

[63] It should be noted that Nickles' reading is not the common way to interpret this example. Typically classical mechanics is said to be reduced to the special theory of relativity (and not the other way round) in the sense that the former is a special case of the latter.

[64] The relation between classical genetics and molecular biology can also be reconstructed as a case of successional reduction. Thus, this example shows that successional and interlevel reduction often cannot be told apart as neatly as a philosopher might wish.

[65] Accordingly, Wimsatt named interlevel reduction "explanatory reduction" (1976a, 677).

does not show that the conceptual distinction between successional and interlevel reduction is misleading in the context of theory reduction. However, Wimsatt's claim indicates an interesting possibility, namely that of also applying the difference between successional and interlevel reduction to the *context of explanations*. What *interlevel reduction* with regard to explanations might be is not difficult to imagine because this is the kind of reduction current discussions about explanatory reduction are focused on. Interlevel reduction pertains either to the reductive relation between a higher-level and a lower-level explanation (at the same time) or between a higher-level explanandum and a lower level explanans of the same explanation.[66] What *successional explanatory* reduction may be is less clear. Analogous to successional theory reduction it could be interpreted as a relation between a historically earlier and a historically later explanation of the same phenomenon. Hence, questions about successional explanatory reduction are related to the development of explanations in one domain across successive historical periods of scientific change.[67]

It should have become apparent from my previous remarks that my analysis of explanatory reduction is primarily concerned with interlevel reduction, not with successional reduction. With respect to this decision one may raise the following concern: to focus on interlevel reduction implies to mistakenly treat explanations as static elements, to take into account only the *context of justification*, and to neglect the process of how an explanation is developed over time (*context of discovery*). In my view there is a persuasive reply to this objection. To focus on interlevel reduction does not commit one to overlook the fact that explanations are far away from being unchanging elements in biological practice. Rather, one can admit that explanations frequently are developed in a stepwise fashion and continuously may be further refined. This is especially true if one focuses on the analysis of individual reductive explanations, as I do. However, this does not amount to considering the issue of successional reduction. Even if one takes into account the context of discovery in analyzing reductive explanations (e.g. by relating the features of reductive explanations to the application of certain reductive methods; see Chapter V) one still conceives of reduction as a relation between (maybe changing) higher-level and lower-level descriptions and not as a relation between a historically earlier explanation and its successor.

To conclude, the third lesson that can be drawn from the previous debate is that one should not only keep apart ontological and epistemic reduction, but also distinguish between three types of epistemic reduction, namely between theory reduction, methodological reduction, and explanatory reduction. Furthermore, it is important to tell apart cases of successional reduction from cases of interlevel reduction. All in all, this section served to introduce these different kinds of epistemic reduction and the major ideas that are linked to them. In addition, I provided grounds for why it is most promising to analyze

---

[66] These two kinds of interlevel reduction correspond to the two types of explanatory reduction characterized in Chapter III.

[67] This might be similar to what Rosenberg (2006) has in mind (see Chapter III, Section 1).

epistemic reduction in biology by focusing on (interlevel) explanatory reduction, and not on methodological reduction or on theory reduction. I argued that my choice to develop an account of explanatory reduction (and not of methodological reduction) is not only due to my personal interests. Rather, reductive explanations are the better targets of philosophical analysis because explanations are, in general, more constrained and less discipline specific than methods. In the next section I complete this argumentation by pointing out why one should back away from regarding reduction as a formal relation between theories.

4 Fourth Lesson: It Is Time to Move beyond Nagelian Reduction

Let me come back to the basic question of this book, namely, What is reduction in actual biological practice? On my way to find an answer to this question this chapter introduces the previous debate about reduction(ism) in philosophy of biology by identifying the main lessons one should learn from it. These lessons at the same time delineate and justify the path I am going to take in answering the basic question: in Section 2.3 I argued that my analysis of reduction focuses on epistemic rather than on ontological issues since cases of epistemic reduction are what can actually be found in biological research practice and what is of concern to biologists. In Section 3 I distinguished three types of epistemic reduction, namely theory, methodological, and explanatory reduction. I argued that reductive explanations are a more promising target of a philosophical analysis than reductive methods because they are more constrained and less discipline specific. What remains for the last section of this chapter is to show why it is time to move beyond Nagelian reduction and to turn our attention to the search for an adequate model of explanatory reduction in biology.

4.1 Clarifying the Object of Criticism

In order to demonstrate the inadequacy of Nagel's understanding of epistemic reduction with regard to biology I do not simply want to echo the old criticism that has been put forward against Nagel's classical model of theory reduction to reveal its general problems and its inapplicability to biology (see Section 3.1). Rather, I abstract from many of the details of Nagel's approach that have turned out to be highly problematic and focus my critique on Nagel's general way of thinking about epistemic reduction. That is, I admit that a proponent of Nagel's model may improve it in three respects:

First, according to Nagel the relation of reduction holds between theories, which he conceives as systems of statements, containing law statements and being formalized in first order logic (Nagel 1961, Chapter 5 and 6; see also Giere 1988, Chapter 3).[68] Nagel's particular view of the relata of reduction encounters a serious objection, namely that the relata of reduction he envisions do not exist in biological practice. Neither do there exist

---

[68] This view is referred to as the syntactic conception of theories.

strict laws in biology nor do biological theories satisfy the demands of the syntactic view of theories. In order to cope with this criticism, proponents of the Nagelian account can choose one of two options: on the one hand, they can give up the notion of a strict law and adopt a more moderate account of what a scientific law is. For instance, they might argue that biological laws are so called ceteris paribus laws (e.g. Lange 2000) or adopt the concept of a "pragmatic law" (Mitchell 1997, 2003). This would allow them to claim that there exist genuine laws in biology and, thus, to argue that the relata for Nagelian reduction, namely theories as sets of law statements, are available. On the other hand, one can counter the critique that Nagel's envisioned relata of reduction do not exist by adjusting the notion of a scientific theory. For instance, one might simply abandon Nagel's claim that theories must consist of law statements. Instead, one could allow each general statement formulated in first order logic to function as relatum of reduction.[69] In principle, it is possible to go even further and to abandon the "syntactic view" (Suppe 2000, 102; see also Thompson 1989) of theories and with it the requirement that theories must be formulated in first-order logic. However, as I will argue in the following section (see Section 4.2), on closer inspection it turns out that this modification leads the Nagelian model too far away from its core ideas and, thus, cannot be accepted.

Second, Nagel's model encounters another objection, namely that the reduced theory may contain empirically false statements and, in this case, cannot be deduced from the reducing theory. In order to address this key drawback one could accept the changes of the Nagelian model Schaffner (1967, 1969, 1993) made in his general "General Reduction-Replacement (GRR) Model" (1993, 429). In line with Schaffner one could claim that an account of theory reduction also captures cases in which not the original theories themselves, but rather corrected versions of the reduced theories are derived from the reducing theories.

Third, many subsequent discussions about Nagel's model have turned on the nature of bridge statements, which are needed in cases of heterogeneous reduction. A frequently raised objection was that since many biological types are multiply realized no bridge statements that express identities between biological types and, for example, molecular types can be found (see Section 3.1). A possibility to sidestep this problem is to abandon the strong claim that bridge statements must be factual claims that express identity relations – a thesis that is typically associated with Nagel's model.[70] Even if the bridge statements are taken to be factual claims, and not, for instance, mere stipulations/conventions, one could argue that it is still left open which ontological relation they express (e.g. mere correlations, necessary nomic connections, constitutional relations, identity relations, etc.; see also Dizadji-Bahmani, et al. 2010, 403f).

---

[69] However, this option turns out to be problematic since Nagel regards reduction as a special case of explanation, which in turn presupposes the availability of law-like generalizations (at least if one endorses the D-N model, as Nagel did).

[70] Though it is doubtful whether Nagel holds this strong view; see 1961, 354-358.

If a defender of the Nagelian account relinquishes all these problematic assumptions (i.e. that theories consist of strict law statements, that theories must not be corrected before derivation, and that bridge statements express identities), what remains is Nagel's general way of thinking about epistemic reduction, which can be characterized by the two core assumptions introduced in Section 3.1:

(1) the adequate units of the relation of reduction are *theories* and
(2) the relation of reduction is a relation of *logical derivation.*

My claim in this section is that even this very moderate, thin version of the Nagelian account of reduction is flawed. In what follows I reveal why it is inadequate to think about epistemic reduction in biology in terms of theories and the logical relations between them (Section 4.3). The general line of my argument will be that a formal model of theory reduction neither captures the most important cases of epistemic reduction in biology, nor accounts for the diversity of reductive reasoning strategies present in current biological research practice. This leaves us with an account of epistemic reduction that has at least a very *restricted range of application* in biology and that provides us with a *misleading picture* of what epistemic reduction in biology is. But beforehand, I need to argue for why one cannot defend Nagel's approach by abandoning the syntactic view of theories (Section 4.2).

4.2 Why Not Abandon the Syntactic View of Theories?

In the previous section I have outlined the possibility to improve Nagel's model of theory reduction by abandoning the "syntactic" or "received view" (Suppe 2000, 102; see also Thompson 1989) of theories and with it the requirement that theories must be formulated in first-order logic. Instead, one could argue for a "semantic view" (Suppe 1977, 1989; Lloyd 1988) of theories, according to which theories are families or sets of models that meet specific set-theoretic conditions.[71] However, I have also argued that that this is *not an option* for a proponent of a Nagelian account because it results in a model of theory reduction that is not "Nagelian" anymore. In this section I elaborate on this argument.

Let us begin with considering the opposite point of view. The possibility of improving Nagel's model by abandoning the syntactic view is one reason why Dizadji-Bahmani, Frigg, and Hartmann (2010) want to convince us not to be afraid of Nagelian reduction anymore. In their view, the syntactic view of theories is "unnecessary" to get Nagel's account "off the ground". We can replace first order logic "with any formal system that is strong enough to do what we need it to do" (2010, 403).

---

[71] Many philosophers of biology have embraced this *semantic view* of theories, especially with respect to evolutionary biology. See, for instance, Beatty 1981; Lloyd 1988; Thompson 1989; Sloep/van der Steen 1987; as well as the different responses to the paper from Sloep/van der Steen in Biology and Philosophy Vol. 2, No. 1.

Likewise, Bickle (1998, 2003) clings to the view that reduction is a relation between theories but argues for a semantic conception of theories. Based on Hooker's (1981) approach to reduction Bickle formulates his "new-wave account of intertheoretic reduction" (1998, 23) according to which the reduction of one theory T_R to another T_B requires the construction of an "image I_B of the set-theoretic structure of models of the reduced theory T_R within the set comprising reducing theory T_B" (2003, 27). The details of Bickle's "semantic" account of intertheoretic reduction are complex. However, what matters for my purposes is that Bickle explicitly contrasts his approach with the Nagelian idea of "characterizing intertheoretic reduction in terms of syntactic derivations" (2003, 27).

This suggests that we can only stick to the thesis that reduction is a relation between theories *and* adopt the semantic view of theories if we abandon the second core assumption of Nagel's account (see Section 4.1) that the relation of reduction is a relation of *logical derivation*. However, it is important to note that this is not true for *all* versions of the semantic conception of theories. According to some notions of a model, it is possible to state that one set of models is logically derived from another set of models (in this sense, e.g., Galilei's theories of motion can be derived from Newton's theories of mechanics). But according to several other versions of the semantic conception of theories, especially according to those that can be found in philosophy of biology, the relation between two theories as sets of models satisfying certain set-theoretic conditions is not one of logical derivation, but rather one of "isomorphism" (e.g. van Fraassen 1980, 46; Lloyd 1988, 14) or one of "analogy" (e.g. Bickle 1998, 32). Hence, if one adopts a version of the semantic conception and a concept of model that accounts for actual cases of theories and models in biological practice *and* if one wants to stick to the claim that reduction is a relation between theories, it seems as if one must abandon Nagel's central assumption that reduction is a relation of logical derivation. But this takes us so far away from the core ideas of Nagel's model of theory reduction that the resultant view of epistemic reduction cannot be called "Nagelian" anymore.

4.3 The Inadequacy of Nagel's Model to Biology

In this section I present what I think are the three most important reasons why Nagel's general way of thinking about epistemic reduction in terms of theories and their logical relations is inadequate with regard to biology. To put it in a nutshell, first, biological research practice shows that, in general, theories are not the only (and perhaps not the most) important element of scientific practice. Second, it reveals that for reduction, in particular, theories are only peripherally important since the most crucial and frequently occurring cases of epistemic reduction rarely involve fully explicated theories. Third, the logical empiricist's focus on formal issues neglects substantive issues, which are important for a proper understanding of epistemic reduction in biology.

4.3.1 Theories Are Not the Whole Story

As I have just argued, even a proponent of a moderate version of Nagel's approach needs to stick to the syntactic view of theories. Accordingly, he is exposed to all the criticism that has been put forward against this conception. These objections can be summarized as follows: First, with its focus on the formal features of theories the syntactic view fails to capture what biological theories in fact are (i.e. the so called "theories in the wild" Craver 2002a, 65). For example, it does not account for the diversity of representations of theories biologists actually use and which are neither restricted to first order logical predicates nor to linguistic representations at all (see e.g. Laura Perini's work on the importance of diagrams in biology; Perini 2005, forthcoming). Second, the syntactic conception focuses on already established, static theories (context of justification) and lacks an account of the dynamics of biological theories, that is, of how they are developed over time (context of discovery) (Darden 1991; Lloyd 1988). Third, the syntactic view overestimates the role of full-established theories by ignoring the important roles other epistemic units (such as models, descriptions of mechanisms, fragments of theories, etc.) play in explanation, prediction, discovery, and manipulation in biology.

The latter criticism can be expanded to a more general objection that applies to any account of the biological sciences that exclusively focuses on the significance of theories – regardless of whether this account is based not on a syntactic view of theories or on the alternative semantic view. First, often not fully explicated theories as a whole, but rather fragments of theories, individual models (as opposed to entire sets of models), and descriptions of particular mechanisms[72] play important roles in explanation, prediction, discovery, and manipulation. Second, in biology there exist several epistemic units that seem to be relatively independent from theories, but that, nevertheless, are crucial for the successful functioning of the biological sciences. Examples include explanatory and investigative strategies (Waters 2008), semi-empirical rules (Sarkar 1992), and mechanistic models (Wimsatt 1976a; Machamer, et al. 2000). Finally, some authors have argued that scientific models, in general, are better conceived as being independent from theories, rather than being constitutive of them (e.g. Morgan/Morrison 1999). All this suggests that a one-sided focus on theories as the only or the most important epistemic units in biology is inadequate.

4.3.2 Almost a Consensus: Nagel's Model Fails

The minor significance of theories to biological practice is particularly apparent in the context of reduction. Briefly speaking, the paradigmatic and most important cases of epistemic reductions that can actually be found in biological research practice are not theory reductions, but rather reductive explanations and reductive methods. Nagel's model just

---

[72] If mechanistic models are understood as being parts of theories (see e.g. Craver 2002a).

does not seem to fit the reality of biology since it does not account for the real cases of epistemic reductions.

The actual research literature in the life sciences clearly supports this thesis. Reductive relations between theories (or even between entire disciplines) are either no subject of discussion at all (Beresford 2010, 721), or theory reduction is explicitly dismissed as being unimportant to biological practice, as the following quotation illustrates:

> Exploring the epistemic relationships between different disciplines [or theories] might be grist in the mill for a philosopher of science but does not seem a particularly fruitful endeavor for a working scientist. (Fang/Casadevall 2011, 1401)[73]

Even more convincing than these self-reports of scientists, is the fact that in the biological literature almost all discussions about reduction, reductionism, and related topics concern explanatory and methodological issues, but not the logical relation between theories.

The insight that the Nagelian account fails to capture the important cases of epistemic reduction in biology is not new (see also Section 3.1). Nowadays, the majority of philosophers of biology hold this view.[74] For instance, Brandon states that "[t]heory reduction… has little or nothing to do with actual scientific practice" (1996, 180). Likewise, Peter Godfrey-Smith claims that "[w]e would probably be better off without it" (2008, 70). And Dupré concludes:

> Within the philosophy of biology, something that has surely received the status of a *consensus* is that no such derivations are plausible. […] If the question of reductionism were merely a question of whether all of biology could be derived from the laws of physics, then we could confidently assert that the issue had been resolved. (2009, 33; my emphasis)

These days even Rosenberg, who defended Nagel's approach for a long time (1985, 110; 1994, 22), demands that "the question of what reductionism was in the postpositivist past" be replaced by "the question of what reductionism is now" (2006, 40). In his latest paper (2006) even Schaffner, the originator of the most important refinement of Nagel's account (i.e. the GRR model), acknowledges the significance of "fragmentary patchy explanations" (2006, 378) and "partial reductions" (2006, 385). Schaffner's main thesis in this paper is that robust reductions of one theory or branch of science by another are "largely a myth" (2006, 377). The kind of reductions one in fact encounters in biology are not "sweeping reductions" (2006, 378), but rather fragmentary patchy explanations, which are only "creeping" (2006, 379), partial reductions. At first sight, this looks as if Schaffner wants to drop his GRR model. But under closer inspection one notices that Schaffner still regards the GRR model as an "ideal" (2006, 384) of what a complete reduction would look like. He admits (as he has done

---

[73] A similar conclusion is reached by Mayr 1988, 475.

[74] In philosophy of mind the view that Nagel's account is inadequate to scientific practice and, thus, needs to be abandoned is far away from being a consensus yet. However, some philosophers of mind have started to put forward this idea (e.g. Kim 1999, Gillett 2007).

earlier; e.g. 1974a; 1993, Section 9.7) that in actual biological practice this ideal is neither "*directive*" nor a "fully accurate summary of the *results* of [molecular biological research]… programs" (2006, 383). But the fact that Schaffner refers to reductive, causal mechanical explanations as mere "*partial* reductions" and "reductions of the *creeping* sort" (2006, 397) strongly suggests that he sticks to theory reductions as the ideal (or, at least, as a "secret hope or end" Wimsatt 1976a, 685). However, Schaffner is one of the very few philosophers of biology who resist the view that Nagel's model is inadequate to real biological practice.

4.3.3 Reasons for the Failure

Let us now turn to the details of why even a moderate version of Nagel's model fails to capture the paradigmatic cases of epistemic reductions that occur in biological practice. In what follows I identify three respects in which Nagel's account does not capture what epistemic reduction in biology really is.[75]

First, with its focus on the relation between theories or entire disciplines the Nagelian model accounts only for *global* and *complete* cases of epistemic reductions. By contrast, the reductions that can be found in biological practice are mostly *local* and *partial* reductions (e.g. Wimsatt 2006a, 448; Wimsatt/Sarkar 2006, 697; Schaffner 2006, 397f). For instance, reductive explanations typically do not span all phenomena that are addressed by a theory or that are studied in a discipline. Rather, they are local in the sense that they explain only small range types of phenomena (i.e. generalizations of small scope) like bacterial chemotaxis or even singular phenomena like the mutated phenotype of a particular *Drosophila* fly. Furthermore, reductive explanations can be called partial since they often cite factors that are located on higher levels than the fundamental level of molecules and genes (but nevertheless on lower levels; see Chapter V, Section 2) and since some of them encompass black boxes that are not filled in yet (Machamer, et al., 18; Schaffner 2006, 397). Also the reductive methods that are applied in the biological sciences are not global stances (see Section 3.2.2), but rather mid-range research strategies or heuristics that are frequently "re-tuned, re-modulated, [and] re-contextualized" (Wimsatt 2007, 10) in order to be suitable to the particular research interests of a certain field.

Second, the Nagelian model of theory reduction has been closely connected with the so called *"layer-cake" picture* of science (Oppenheim/Putnam 1958; see also Chapter V, Section 1.3.1).[76] According to this picture, biology is organized into separate sciences, each of which aims to discern the laws governing the behavior of the objects at a particular level of organization, like the level of ecosystems, populations, organisms, organs, cells, or molecules. Given this framework, reduction is assumed to be the deduction of a higher-level

---

[75] In order not to expand this critical discussion I invoke just a few concrete examples of actual cases of epistemic reduction in this section. However, in Chapter V I analyze several of these examples, each of which could be adduced as empirical evidence for my theses here.

[76] Granted, one can question whether Nagel's model is committed to the layer-cake view of science (see e.g. Steel 2004, 60).

theory from a lower-level one.[77] What is problematic with this layer-cake picture of biology is that it seems to reflect the idealizations of philosophers, but fails to correspond to the actual science. Neither biological fields or theories nor biological explanations are confined exclusively to one layer of a cake (i.e. one level of organization). There are no unconnected tiers of theoretical discourse, but rather many kinds of overlapping, interrelations, and integrations between different biological fields. Moreover, biological explanations typically exhibit a *multilevel* character (e.g. Craver 2007a, 9-16; Mitchell 2009, 109-115) – and this holds also for many reductive explanations in biology. In addition, in light of the actual explanatory practice in biology the layer-cake assumption that the lowest levels of organization always have explanatory primacy cannot be sustained (see also Chapter IV, Section 2.2). Several biological phenomena cannot be reductively explained and even reductive explanations are not restricted to fundamental level explanations (see Chapter V, Section 2.3.1).

Finally, the overall critique that lies behind most of these objections is that in the tradition of logical empiricism Nagel's model of epistemic reduction centers on *formal* issues (like the logical relations between sentences formalized in first order logic) and, thereby, neglects significant "*substantive* issues" (Sarkar 1998, 19; my emphasis). For instance, Wimsatt convincingly argues that the features of epistemic reduction need to be analyzed in terms of their "functioning in efficiently promoting the aims of science [e.g. explanation]" (1976a, 700) and that it is highly questionable that a formal model accomplishes this (1976a, 673-675). Another point is that models of theory reduction are relatively free of ontological commitments.[78] However, biological practice suggests that knowledge about ontological issues, like part-whole relations and hierarchies involving different levels of organization, are crucial for performing epistemic reductions. Hence, a philosophical account of epistemic reduction should include these substantive functional and ontological issues. As the title of this book ("An *ontic* account…") suggests ontological issues will occupy center stage in my own account of explanatory reduction (see Chapter V, Section 6).

In sum, even if one concedes that the proponents of the Nagelian model of theory reduction can handle several problems that have been raised in the past (Section 4.1 and 4.2), Nagel's general way of thinking about epistemic reduction in terms of theories and their logical relations still remains inadequate with regard to actual biology. It is inadequate because it exclusively considers theories, which are neither the only nor the most important epistemic units in biology (Section 4.3.1) and which are particularly insignificant in the context of epistemic reduction. More precisely, even the most moderate version of Nagel's model fails to capture what epistemic reduction in biological practice really is since it accounts only for global and complete cases of epistemic reductions, since it is closely

---

[77] This layer-cake view of reduction has also been named the "standard account of reduction" (Kincaid 1990, 576).

[78] Unless bridge statements are interpreted as expressing identity relations between types (see Section 2.2).

connected to the fallacious layer-cake picture of biology, and since it neglects significant substantive issues concerning epistemic reduction (Section 4.3.2 and 4.3.3). Thus, the fourth lesson we should learn from the previous debate is that it is time to move beyond Nagelian reduction and to shift the attention from theory reduction to reductive explanations and from formal to substantive issues.[79]

5 Interim Conclusion

The goal of this chapter was to introduce the previous debate about reductionism in the philosophy of biology. However, instead of giving a mere diachronic overview about the debate I identified what I conceive as the most crucial lessons one should learn from this debate. In doing so, I introduced important concepts and distinctions, and showed the reader the path I will run in the remaining part of this book.

My main thesis in this chapter is that there are four lessons one should draw from the previous debate: *first*, before you discuss whether reductionism or antireductionism is true you should seek to understand what reduction is (Section 1). In other words, you should recognize that the project of understanding reduction is prior to the project of disputing reductionism. Any convincing argumentation in favor of a particular version of reductionism or antireductionism must be based on a clear and precise understanding of what reduction is. From this it follows that my project of developing an account of explanatory reduction in biology has the potential to enhance disputes about explanatory reductionism. We can much better argue about the feasibility, the adequacy and, in general, the scope of explaining biological phenomena reductively if we are very clear about what it means to provide a reductive explanation of biological phenomena. However, even if it is clear which kind of reduction is under discussion, there is still a bewildering variety of positions available that a (anti-)reductionist can take up. In Section 1 I proposed four criteria to classify different versions of reductionism: first, is reductionism a claim about reduction in practice or about reductions that are achievable only in principle; second, is reductionism eliminative or retentive; third, what is the scope of reductionism; and fourth, what is the reducing realm (i.e. locus) that reductionism assumes?

The *second lesson* that can be learned from the previous debate is twofold: on the one hand, if you aim at understanding what reduction in actual biological practice is, you should be aware of the fact that epistemic issues (and not ontological issues) are the ones that matter most (Section 2.3). But yet, you should, on the other hand, recognize the respects in which epistemic and ontological issues are interrelated, and you should keep an eye on that epistemic issues are not confused with ontological ones, as it is often the case (Section 2.2).

---

[79] In line with this, Sarkar claims that "models of explanatory reduction… often seem best suited to capture the flavor of actual scientific reductions." (1992, 175)

The latter task requires that philosophers of biology not dismiss the ontological questions in such as hasty a manner as they actually do. There is more to be said about ontology than "We're all ontological reductionists. Case closed". This is why, in Section 2.1.1, I specified the kind of ontological reductionism that is (often implicitly) taken for granted in the debate about reduction(ism) in philosophy of biology. I did this by borrowing some concepts and distinctions from philosophy of mind. The result of this analysis is that the type of ontological reductionism that constitutes the implicit consensus in philosophy of biology is token physicalism (whereas supervenience physicalism alone is regarded as non-reductionistic). Furthermore, it is left open whether the reduction relation must be specified as a relation of identity or as a relation of constitution/localization (and what the difference between the two is). Thus, contrary to the situation in philosophy of mind, the notion of ontological reduction that is applied by philosophers of biology is not confined to the identity relation between types or properties.

After having clarified the notion of ontological reduction (Section 2.1.1) and of epistemic reduction (Section 2.1.2) I then turned to the question which relations exist between the two issues (Section 2.2). The first answer I gave was negative. In Section 2.2.1 I rejected Rosenberg's thesis that you cannot have ontological reduction without at the same time having explanatory reduction and Brandon's assumption that ontological and explanatory reduction fall together. I showed that both authors derive their claims about the very tight connection between ontological and explanatory reduction from an implausible view about scientific explanation. Even if we accept that "we live in one world" and that any particular biological entity is constituted by nothing but physical entities (token physicalism), we are still not committed to the reductionist claim that these physical entities represent a *privileged* level of description. However, this is not to say that there exist no connections between ontological and explanatory reduction at all (Section 2.2.2). As my own analysis of explanatory reduction (see Chapter V) will clearly display, the availability of an adequate reductive explanation of the behavior of a particular system presupposes that certain relations and facts exist in the world (e.g. compositional relations, spatial localizations, etc.).

Finally, I argued that philosophers of biology have a persuasive reason for why they have focused so much on epistemological issues concerning reduction during the last 50 years (Section 2.3). The biological literature (which will be analyzed in more detail in Chapter V) shows that epistemic reduction, not ontological reduction, is of central concern to biologists. More precisely, they are primarily concerned with methodological and explanatory issues.

The *third lesson* that can be learned from the previous debate is that you should not only keep apart ontological and epistemic reduction, but also distinguish between three types of epistemic reduction, namely between theory reduction (Section 3.1), methodological reduction (Section 3.2), and explanatory reduction (Section 3.3). Furthermore, it is important to tell apart cases of successional reduction from cases of interlevel reduction (Section 3.4). In short, the lesson is that before you discuss about reduction(ism) you need to specify which kind of reduction you are talking about. In general, I pursued two goals in this section.

On the one hand, the task was to introduce the different kinds of epistemic reduction and to discuss the major ideas that are linked to them. In doing so, I also broke fresh ground, most notably in specifying the concept of methodological reduction. In Section 3.2.1 I argued that common formulations of methodological reductionism do not sufficiently specify. I undertook first steps to fill this gap (Section 3.2.2) by characterizing reductive methods as middle range reductive heuristics or investigative strategies and by introducing the reductive method per se, namely decomposition (further reductive methods will be specified in Chapter V). Moreover, I revealed that the usage of reductive methods can but need not result in reductive explanations (Section 3.2.3). For instance, if the application of reductive methods fails or if they are employed together with non-reductive methods they commonly do not give rise to (adequate) reductive explanations.

On the other hand, my aim was to adduce reasons for why it is most promising to analyze epistemic reduction in biology by focusing on (interlevel) explanatory reduction, and not on methodological reduction or on theory reduction. In Section 3.2.4 I stated that methodological issues concerning reduction are significant and do play some role in my analysis (see Chapter V). However, I also emphasized that my choice to develop an account of *explanatory* reduction and not of methodological reduction is not only due to my personal interests. In addition, I claimed that reductive explanations are the better targets of philosophical analysis because explanations are, in general, more constrained and less discipline specific than methods.

Philosophical analyses of epistemic reduction have concentrated on reduction as a relation between theories for a long time. In my view, the *fourth lesson* one should draw from the previous debate is that it is time to move beyond the limitations of Nagel's classical model of reduction and, instead, to seek an adequate model of explanatory reduction in biology. My argumentative strategy was to show that, even if one admits that Nagel's original model can be improved in several respects (Section 4.1 and 4.2), his general way of thinking about epistemic reduction in terms of theories and their logical relations will remain inadequate with respect to actual biological practice. It is inadequate because it exclusively considers theories, which are neither the only nor the most important epistemic units in biology (Section 4.3.1) and which are particularly insignificant in the context of epistemic reduction. Even a refined version of Nagel's model fails to capture what epistemic reduction in biological practice really is since it accounts only for global and complete cases of epistemic reductions, since it is closely connected to the fallacious layer-cake picture of biology, and since it neglects significant substantive issues about epistemic reduction (Section 4.3.2 and 4.3.3).

III Two Perspectives on Explanatory Reduction

"Reductionism is the thesis that biological… explanations… need to be grounded in molecular biology and ultimately physical science, for it is only by doing so that they can be improved, corrected, strengthened, made more accurate and more adequate, and completed."

(Alex Rosenberg 2006, 4)

"[W]e must ask what substantive criteria distinguish reductionist explanations from other forms of explanation. Reductionism then becomes the empirical thesis that explanations in a particular discipline satisfy those criteria."

(Sahotra Sarkar 2008, 427)

Content

In the preceding chapters I revealed the methodological preliminaries of my analysis of reduction in biology (Chapter I), I introduced the previous reductionism debate and, thereby, pointed to the direction in which my analysis will proceed (Chapter II). Thus, the methodological procedure by which I develop my account of reduction and the kind of account I seek are now clearer. I aim to develop an account of *explanatory* reduction – not of ontological, theory, or methodological reduction. More specifically, my goal is to understand what explanatory reduction in contemporary biological practice actually is (reduction *in practice*). In order to do this I analyze the features of paradigmatic and important examples of reductive explanations from current biology as well as their strengths and limitations (i.e. the conditions of their adequacy). However, before this analysis can start (Chapter V), we should take a look at the recent debate about explanatory reduction in philosophy of biology and review the proposals that have been made so far (this chapter). Furthermore, we should examine the notion of explanation that is presupposed in my analysis (Chapter IV).

Hence, the task of this chapter is set. It is to critically discuss the different notions of explanatory reduction that have been proposed (or implicitly assumed) so far. The goal of this critical discussion is, on the one hand, to identify adequate and fruitful insights that will be taken up in my own analysis of explanatory reduction. On the other hand, my aim is to sort out those ideas about explanatory reduction that prove to be unconvincing. All in all, this chapter constitutes a first step towards an adequate understanding of what explanatory reduction in actual biological practice is.

Before I start examining the different positions let me add two preliminary remarks. First, as I have already pointed out in the previous chapter, the topic of reduction has always been closely linked to the topic of explanation. That is, the difference between the early stages of the debate, in which thinking about reduction in terms of theories and their logical relations prevailed, and the current stage, in which almost everyone has abandoned Nagel's model, is *not* that explanatory reduction was the subject of discussion *only* in the latter stage. Rather, the issue of reductive explanations has been addressed in both stages, but the focus of the discussions has been broadened and the kinds of questions asked have changed.

Let us dwell on this point a bit. Nagel himself regarded theory reduction as a special case of explanation and devoted an entire chapter to "Mechanistic Explanation and Organismic Biology" (1961, 398-446), in which he discusses the explanatory autonomy of biology. According to this, when Nagel's model of theory reduction was applied to biology one controversially disputed question was whether a derivation of laws of gene transmission from laws of molecular biology explains why the laws of gene transmission hold (Kitcher 1984, 339, 347-351). But in the framework of Nagel's approach only a special kind of reductive explanation was considered, namely the explanation of a reduced theory by a reducing theory. Furthermore, in accordance with the D-N model of explanation, the logical features of explanations were conceived as being the important ones. These two aspects radically changed the more philosophers of biology treated Nagel's model of theory reduction as inadequate with respect to biology. Apart from reductive explanations of generalizations with a quite broad scope (like theories), proponents of models of explanatory reduction also take into account reductive explanations of token phenomena and of small range types of phenomena (which are by far more common in biological practice; see also Chapter II, Section 4.3.3). That is, most of them consider the entire diversity of reductive explanations in biology. In addition, they either defend an alternative account of explanation (i.e. the causal-mechanistic account; e.g. Wimsatt 1976a) or try to develop an account of explanatory reduction that is as neutral as possible with regard to a certain view of explanation (e.g. Sarkar 1998; see also Chapter IV). However, the notions of explanatory reduction contemporary philosophers of biology put forward differ in the extent they depart from the Nagelian understanding of explanatory reduction. Some depart further than others. This leads us to the second preliminary remark.

Second, it seems to me that in the current debate about reduction in biology basically two different perspectives on explanatory reduction can be distinguished. The first perspective is the understanding of explanatory reduction that implicitly underlies Rosenberg's defense of "Darwinian Reductionism" (2006). He treats explanatory reduction as a relation between a higher-level and a lower-level explanation of the same phenomenon. Contrary to this approach, Sarkar (1992, 1998, 2005), and more recently Hüttemann and Love (2011) focus on individual explanations and examine conditions for their reductive character. Important precursors of the second kind of analysis are Kauffman (1970) with his work on part-whole explanations and Wimsatt (1976a and 2007), who was one of the first to stress the importance of reductive explanations to the biological sciences (1976a, 671). As will become apparent in the course of this chapter, Rosenberg's perspective departs less from Nagel's view of explanatory reduction than Sarkar's, Hüttemann's, and Love's perspective does.

My aim in this chapter is to introduce these two perspectives on explanatory reduction, to identify the fruitful insights they comprise, to discuss the problems they encounter, and, finally, to disclose the connections between them. In doing so, I restrict my investigation to the different notions of explanatory reduction they (explicitly or implicitly) endorse. Questions about the correctness of reductionism (e.g. about the plausibility of Rosenberg's defense of explanatory reductionism) will be addressed only insofar as they are necessary for the understanding of the account of explanatory reduction. The result of this critical examination will be that Rosenberg's understanding of explanatory reduction faces many serious problems, whereas Sarkar's, Hüttemann's, and Love's perspective seems to be a much more promising way to go.

I start my review of previous perspectives on explanatory reduction with an examination of Rosenberg's perspective (Section 1). After reconstructing the major aspects of his defense of Darwinian reductionism (Section 1.1), I reconstruct the notion of explanatory reduction he presupposes in his argumentation (Section 1.2). The result is that

Rosenberg treats explanatory reduction as a relation between two specific kinds of explanations of the same phenomenon, but that the exact nature of the process of reduction remains obscure. I conclude by pointing out some serious criticisms Rosenberg's account faces (Section 1.3).

In the second part of this chapter I analyze different versions of what I regard as the second main perspective on explanatory reduction (Section 2). What binds authors like Kauffmann, Wimsatt, Sarkar, Hüttemann, and Love together is that they focus on individual reductive explanations and investigate the constraints of their reductivity. I begin by giving a brief overview of the first insights Kauffman and Wimsatt gained into the reductive character of biological explanations (Section 2.1).

Afterwards, I consider Sarkar's work on reductive explanations in genetics and molecular biology (Section 2.2). I explicate what it means that Sarkar calls his account substantive (Section 2.2.1) and why he emphasizes its neutrality with respect to any account of explanation (Section 2.2.2). Then I focus on the core of his account, namely his three criteria of reductivity (Section 2.2.3). I conclude by revealing the major shortcomings of his analysis (Section 2.2.4).

In Section 2.3 I discuss the most recent work on explanatory reduction in biology, namely Hüttemann's and Love's analysis of three aspects of reductive explanation. After delineating their methodological framework (Section 2.3.1) I scrutinize their three aspects, namely intrinsicality, fundamentality, and temporality (Section 2.3.2), and clarify in which sense these aspects constrain the reductive character of a biological part-whole explanation (Section 2.3.3). Finally, I disclose the difficulties their analysis encounters (Section 2.3.4).

In the conclusion (Section 3) I summarize what I take to be the most important insights and shortcomings of the previous perspectives on explanatory reduction. I argue that the second perspective, which focuses on individual reductive explanations and analyzes the conditions under which explanations succeed and fail to be reductive, seems to be much more promising that Rosenberg's perspective that treats reduction as a relation between two different kinds of explanations of the same phenomenon.

1 First Perspective: Reduction as a Relation between Two Explanations

To be clear right from the beginning: Rosenberg is not primarily interested in my project of seeking an understanding of what epistemic reduction in biology really is. He pursues the other project of disputing about the correctness of reductionism (see Chapter II, Section 1). Rosenberg takes up the position most philosophers of biology vehemently reject, namely explanatory reductionism (or "Darwinian reductionism", as he calls it). He does not aim at providing a detailed and explicit description of what he means by explanatory reduction.[1]

---

[1] Also Weber evaluates this as a drawback when he claims that Rosenberg's book "suffers a bit from the lack of an explicit account of reduction" (2008, 151).

Thus, one of my tasks in this section will be to single out Rosenberg’s few remarks on what explanatory reduction is from his overall argumentation and to try to integrate them to a unified perspective on explanatory reduction.

I start by sketching Rosenberg’s way to reductionism (Section 1.1.1). After that I characterize the version of explanatory reductionism Rosenberg advocates (Section 1.1.2). To do this, I use the classification I introduced in Chapter II, Section 1. Then I give a rough outline of the general line of argumentation by which Rosenberg defends reductionism (Section 1.1.3). These three sections provide the basis on which I can answer my central question, namely what is Rosenberg’s understanding of explanatory reduction (Section 1.2). Finally, I point out several difficulties Rosenberg’s notion of explanatory reduction faces (Section 1.3). This critical discussion will be confined to the problems Rosenberg’s view of explanatory reduction encounters. The many other criticisms one could raise with regard to his defense of Darwinian reductionism will not be taken into account.[2]

1.1 Darwinian Reductionism

1.1.1 Rosenberg’s Way to Reductionism

Rosenberg has been concerned with the topic of reductionism in biology since the early 1980s.[3] From his work one gets the impression that he has always been strongly attracted by the idea that biology can be reduced to molecular biology and ultimately to physics and, thereby, be systematized and unified. Rosenberg’s strongly reductionistic intuitions arise from the impressive advancement that molecular biology has undergone since Watson and Crick discovered the double helix structure of DNA in 1953. Accordingly, in his work Rosenberg spends much time on highlighting this success story of molecular biology and its bright future prospects.

However, Rosenberg has not always been a defender of reductionism in biology. Rather, this is a novelty of his recent position presented in his latest book “Darwinian Reductionism” (2006). What is important to note is that this change has been made possible by his abandonment of Nagel’s account of theory reduction. In his recent work he focuses no longer on the questions whether certain biological theories can be logically derived from molecular biologists’ theories and how the necessary connections between the terms of the reduced and the reducing theory can be established. Instead, he shifts his attention from theories to explanations and defends a special version of explanatory reductionism (see Section 1.1.2).

By contrast, Rosenberg’s earlier work on reductionism is still soaked with Nagel’s view of reduction and with the problems Nagel’s model encounters in biology. But despite

---

[2] To get an overview of these difficulties see, for example, Weber (2008) and Love (2008b).

[3] Besides his three major books “The Structure of Biological Science” (1985), “Instrumental Biology or the Disunity of Science” (1994), and “Darwinian Reductionism” (2006) he has published an immense bulk of papers.

these difficulties, in 1985 Rosenberg is still optimistic. He claims that the accomplishment of theory reduction in biology is just a matter of "patience and industry" (1985, 72). Impressed by the progress of molecular research in the decades before Rosenberg claims that the complexity of the relation between heredity phenomena and their molecular basis does not reveal the *in principle impossibility* of the deduction of Mendel's laws to molecular biology. Since an "omniscient creature in possession of all the relevant facts could effect this deduction" (1985, 110), theory reduction is physically possible. Thus, the obtaining obstacles must be just "practical, instrumental one[s]" (1985, 110) that are due to the weak cognitive capacities of humans and their inability to deal with the complexity of the world.[4]

The proceeding discussion in the 1980s and 1990s revealed more and more problems with applying Nagel's model of reduction to biology (see Chapter II, Section 3.1 and 4). Rosenberg responded to these findings with asserting the *instrumental character* of all biological theorizing (1994). His line of reasoning is less optimistic than nine years before: above the level of molecules nature is very complex. This complexity arises from the fact that natural selection, which has shaped most of the functional features of biological systems, is blind to structural differences with similar functions. Unfortunately, the cognitive powers of humans are limited and do not enable us to deal systematically with nature's complexity. This explains why there are no strict laws in biology of the sort we are familiar with in physics and chemistry and "why the smooth reduction of biological theory to physical theory is not on the cards" (1994, 55). Since Rosenberg refrains from calling into question the unity of science (contrary to Duprè 1993) he adopts an instrumentalistic, anti-realistic interpretation of biological theorizing. That is, he conceives biology as a "body of claims each of which is qualified by an implicit appeal to its usefulness for cognitive agents of our powers" (1994, 54). As implausible as one might find this conclusion, for Rosenberg it follows: "If reductionism is wrong, instrumentalism is right" (1994, 38).

Against this background one might ask: Why did Rosenberg not abandon Nagel's model and adopt an account of reduction that allows him to adhere to reductionism? It is interesting from my point of view that Rosenberg, in fact, sees himself confronted by these two options – that is, *either* change the notion of reduction *or* accept that Nagelian reduction is not to be had in biology, but that he does not hesitate to choose the latter, which he conceives as "more reasonable" (1994, 22). It seems as if the first option is not a real option for Rosenberg at all. Nine years earlier he had already stated that it "sounds suspicious to change the standards of reduction" (1985, 110). This clearly shows that in the 1980s and 1990s the time had not yet come to step outside the Nagelian framework. Only in his recent work does Rosenberg explicitly abandon Nagel's model as irrelevant to current biology:

---

[4] This deprecatory judgment about the limited cognitive powers of humans can also be found in Rosenberg's latest work. But here it appears as a deficient argumentative strategy of the explanatory antireductionist (see 2006, 14f, 36).

> [T]he 'layer-cake' reductionism of postpositivist philosophers of science… [is] irrelevant to the real issue about the relation between functional and molecular biology. […] [T]he question of what reductionism was in the postpositivist past can be replaced by the question of what reductionism is now. […] It is now clear that the question has to be reformulated if it is to make contact with real issues in biology. (2006, 40)

In sum, Rosenberg has undergone a striking development from the 1980s until now: he started with arguing for in-principle theory reductionism in Nagel's sense (1985). About a decade later he accepted the impossibility of reductionism in biology altogether, still adhering to the Nagelian account (1994). Another decade later he switches to the opposite site and becomes a defender of reductionism in biology by abandoning Nagel's model of theory reduction. Instead of thinking about reduction in terms of theories and their logical relations he now treats reduction as a relation between explanations (2006).

1.1.2 What Does a Darwinian Reductionist Claim?

Rosenberg distinguishes reductionism as a metaphysical thesis, as a claim about explanations, and as a research program. Since he regards physicalism as uncontroversial and methodological questions as subordinate, he concentrates his attention on the defense of explanatory reductionism in biology (2006, 4, 25-27). According to Rosenberg, explanatory reductionism is the thesis that:

> explanations in functional biology need… [to] be corrected, completed, or otherwise made more adequate by explanations in terms of molecular biology. (2006, 26)

Two things should be noted concerning this quote. First, in this passage Rosenberg states that biological explanations are made "more adequate" by reducing them to molecular explanations. This sounds as if non-reduced functional biology's explanations are already (a bit) adequate. However, Rosenberg's overall argumentation suggests that he better should have dropped the term 'more' (and he does so in other similar passages). What he actually seems to claim is that biological explanations are adequate *only if* they are molecular explanations or transformed into (respectively, reduced to) molecular explanations. Non-molecular biological explanations are not *less* adequate, but *none* adequate or successful explanations *at all*.[5]

Second, it is not the case that Rosenberg simply wants to remove functional biology from the scientific landscape, so that biology amounts to nothing more than molecular biology. Rather, he insists that his version of reductionism is *non-eliminative*:

> "reductionism does not eschew the employment of concepts, terms, kinds, and taxonomies that characterize phenomena in nonmolecular terms. Reductionism is not eliminativism" (2006, 84).

---

[5] Since Rosenberg also speaks about inadequate and false explanations he does not seem to use the term 'explanation' as a success term.

Rosenberg illustrates this retentive character of reductionism with a joke, whose lesson is the following: heaven belongs to molecular biologists, but there is still room left for functional biologists in the molecular biologists' heaven (2006, 1f). But on closer inspection it becomes apparent that it is far from being clear what this room for functional biologists is, that is, what exactly the non-eliminative character of Rosenberg's explanatory reductionism amounts to. What Rosenberg *does* point out is that functional biology cannot be eliminated in the sense that it is still needed to identify many of the phenomena to be explained. But he also claims that "functional biology's explanantia are always molecular biology's explananda" (2006, 54). This quote suggests that the task of functional biology is merely to identify the phenomena to be explained, but not to explain them. The latter seems to be the exclusive privilege of molecular biology. According to this interpretation functional biology's *concepts* (including the explananda formulated in these terms) would not be eliminated, but its *explanations* (i.e. the explanatia) would be. However, Rosenberg also talks about the "transformation" (2006, 53) of functional biology's explanations into molecular explanations and about the filling in the links in causal chains with molecular details (2006, 46). As opposed to the former interpretation, this suggests that not only functional biology's concepts but also parts of its explanations are preserved in the course of reduction since they constitute the *starting material* for the reduction process (for further details see Sections 1.1.3 and 1.2).

Let us now take up the classification I introduced in Section 1 of the preceding chapter and apply it to Rosenberg's position. There I argued that a certain version of reductionism can be characterized according to four criteria: (1) in principle – in practice, (2) eliminative – retentive, (3) scope, and (4) locus. When we apply these criteria to Rosenberg's Darwinian reductionism we get the following picture: First, Rosenberg's explanatory reductionism is a claim about what reductions are feasible *in principle*. He neither argues that all explanations that are actually developed in practice are molecular explanations, nor does he state that all transformations (i.e. reductions) of functional biology's explanations into molecular explanations are actually feasible in practice. Rather, he claims that explanatory reductions are accomplishable in principle (and that we should aspire to actually implement them). Second, as I have already pointed out above, Rosenberg insists that the kind of reductionism he endorses is *retentive* and does not lead to eliminativism. However, I have also revealed that it remains unclear what exactly it means for functional biology not to be eliminated and which elements are in fact retained during the transformation of functional biology's explanations into molecular explanations. Third, Rosenberg defends a version of reductionism that has a very *broad scope* since it spans over *all* explanations in biology and is not confined to a certain kind of explanation or to the explanations in a specific field. Fourth, Darwinian reductionism seems to presuppose a very *restricted locus* of explanatory reduction, namely the level of molecules. Rosenberg's claim is that all biological phenomena can only be adequately explained in terms of molecules since only molecular explanations are maximally complete explanations. However, as I mentioned before, it is not clear whether Rosenberg allows *exclusively* molecular terms in the explananda or whether a few functional and higher-level terms survive the transformation and enrichment process (see Sections 1.1.3 and 1.2).

All in all, Rosenberg aims high. Even if he presents his Darwinian reductionism as being non-eliminative in character, not many demanding tasks seems to be left over for functional biologists (under which Rosenberg subsumes any biological field except molecular biology). Furthermore, with its broad scope and its restricted locus, Rosenberg chooses a version of explanatory reductionism that could not be more radical (and more difficult to defend): he claims that *any* biological phenomenon can only be adequately explained on a *single* level, namely the molecular level. Let us now take a look at the arguments Rosenberg offers in support of this position.

1.1.3 Rosenberg's Arguments in Favor of Explanatory Reductionism

Before we can understand how Rosenberg tries to vindicate explanatory reductionism in biology we need to take note of some important distinctions Rosenberg presents. The first is the distinction between two parts of biology, namely *molecular* biology and "the rest of the discipline" (2006, 2), which Rosenberg names "*functional* biology" (2006, 25; my emphasis). According to him, functional biology studies phenomena under their functional kind-descriptions (e.g. wing, chloroplast, heart, etc.), whereas molecular biology is concerned with certain classes of organic macromolecules as well as their interactions. Rosenberg admits that this distinction is not entirely satisfactory (e.g. since molecular biology also individuates some kinds functionally). But yet he employs it as a "handy label" (2006, 25) for the two parts of biology whose relationship is disputed by reductionists and antireductionists.

The second distinction is the one between *ultimate* and *proximate* explanations that was introduced to the debate by Mayr (1982). These two kinds of explanations constitute answers to two different kinds of questions that can be asked about biological phenomena, like the eyespots on the wings of buckeye butterflies. Ultimate explanations answer the question of why the trait "spots on wings that resemble owl eyes" was established in populations of *Junonia coenia* over time, that is, of why it was an adaptation for buckeye butterflies (living in a certain environment) to possess spots on their wings that resemble owl eyes. By contrast, proximate explanations provide an answer the question, how individual buckeye butterflies develop the spots on their wings, that is, which genes code for the color pigments that are stored in the upper layers of skin of the wings, how these genes are regulated and expressed, etc. In other words, ultimate explanations cite the distal causes that lead to an event and proximate explanations the proximate causes (for further details see Chapter IV). As Rosenberg emphasizes, biology (in contrast to physics and chemistry) seeks ultimate evolutionary explanations along with proximate explanations (2006, 17).

The third distinction, which is probably the most important one for Rosenberg's argumentation, is the distinction between *how-possible* and *why-necessary* explanations. A how-possible explanation shows how something *could* have happened, "by adducing facts which show that there is, after all, no good reason for supposing it could not have happened" (2006, 43). As opposed to this, a why-necessary explanation reveals that its explanandum *had to* have happened. Although Rosenberg admits that each of these two kinds of explanation will be appropriate to a different inquiry, he quickly calls the reader's attention to the "important asymmetry" (2006, 43) that how-possible explanations motivate the search for why-necessary explanations, but not vice versa. According to Rosenberg, this is due to the fact that the latter are "more complete" and, thus, "closer to the whole story" (2006, 44) than the former.

On the basis of these three distinctions we can now reconstruct Rosenberg's arguments for Darwinian reductionism. To put it briefly, his argumentation runs as follows: Functional biology yields ultimate explanations as well as proximate explanations, which are also implicitly ultimate. Functional biology's ultimate explanations are only incomplete how-possible explanations. Since in the context of advanced biological inquiry only maximally complete explanations are adequate, functional biology's ultimate how-possible explanations need to be turned into (i.e. reduced to) molecular proximate why-necessary explanations. Let us have a more thorough look at the argumentation.

The first step Rosenberg takes seems to lead him further afar from explanatory reductionism. He accepts Theodosius Dobzhansky's dictum "Nothing in Biology Makes Sense Except in the Light of Evolution" (1973) as literally true. In Rosenberg's view this amounts to accepting the thesis that "every proximate explanation in biology is implicitly ultimate" and, thus, includes an "implicit commitment to the theory of natural selection" (2006, 20). Rosenberg holds this thesis since he assumes that in biology (contrary to in physics and chemistry) the classification of objects into types is based on the functional properties of objects. For instance, a wing is a wing because of its function to enable its bearer to fly. And since Rosenberg endorses an etiological account of functions, each function ascription inevitably involves the reference to natural selection. All this seems to lead Rosenberg further afar from a defense of explanatory reductionism since evolutionary explanations traditionally have been regarded as the worst obstacle for reductionism because of the alleged irreducibility of the principle of natural selection. However, Rosenberg claims to overcome even this obstacle by showing that the principle of natural selection is a "nonderived but physically unproblematical law of chemistry" (2006, 200). Thus, for him the concession that all biological explanations implicitly invoke natural selection is not an obstacle to reductionism.

Let us now turn to the core of Rosenberg's defense of explanatory reductionism. It is based on the assumption that the three distinctions introduced above can be easily assigned to one another: The ultimate explanations (i.e. genuine ultimate explanations and proximate explanations that are implicitly ultimate) that are developed in functional biology are incomplete how-possible explanations (2006, 47-49). By contrast, molecular biologists construct proximate explanations that are complete why-necessary explanations. From this Rosenberg follows that reduction must turn the "merely how-possible scenario of the functional ultimate explanation into a why-necessary proximate explanation of a historical pattern" (2006, 46) since only molecular why-necessary explanations are maximally complete and, thus, adequate.

Of course, the picture of these two biological fields that fall apart so neatly seems to be too simple to capture the reality and diversity of the biological sciences. And even if one accepts this picture as being at least approximately true, the successful vindication of explanatory reductionism still relies on a very difficult presupposition, namely the assumption that biological explanations are adequate only if they are *maximally complete* and that they can only be maximally complete on the *molecular level* (see also Chapter II, Section 2.2). In my view, this is the core assumption, on which much of Rosenberg's argumentation hinges. In his words, it reads as follows:

> [R]eductionism holds that there is a full and complete explanation of every biological fact, state, event, process, trend, or generalization, and… this explanation will cite only the interaction of macromolecules. (2006, 12)

There is an obvious objection that can be raised with respect to this assumption. A proponent of what Rosenberg calls the "erotetic account of explanation" (2006, 35) could object that there is no such thing as a complete explanation independent of its context. He could continue that the adequacy of an explanation depends on the question being asked, as well as on the context in which the question is being asked, that is, on the interests and background knowledge of the recipients. Consequently, for instance a proximate explanation is no adequate answer to an ultimate question and a less complete explanation can in some contexts be an entirely adequate explanation. At first sight, Rosenberg seems to agree since he states that "everyone… will grant that there are many explanations… which are accepted as appropriate to the contexts in which they are given" (2006, 13). But it quickly becomes clear that Rosenberg rejects the erotetic account of explanation. He reproaches it for relegating (anti-)reductionism to "a claim about biologists, not about biology" (2006, 36). Rosenberg insists that in the context of advanced biological inquiry "there is such thing as a complete and correct explanation *independent* of contexts of inquirers' questions" (2006, 44; my emphasis).[6] According to him, this maximally complete and correct explanation must be a proximate why-necessary explanation in macromolecular terms and it is the only kind of explanation that is adequate (in the context of advanced biological inquiry). As a result, all explanations in functional biology (which are only ultimate

---

[6] There is a possible objection to Rosenberg's claim that maximally complete molecular explanations are always better or the only adequate explanations. For instance, Elliott Sober (1999) has argued that depth is only one of two possible virtues of explanations (the other is generality). This point is elaborated in Chapter IV.

how-possible explanations) need to be made adequate by converting them into (i.e. reducing them to) molecular proximate why-necessary explanation.[7]

1.2 Rosenberg's Notion of Explanatory Reduction

On the basis of this overview of Rosenberg's argumentation we can now turn to the central question, namely, What is his view of explanatory reduction?[8] Rosenberg treats explanatory reduction as a certain relation that exists between different types of explanation of the same phenomenon. The *relata of reduction* are ultimate how-possible explanations from functional biology that are reduced to proximate why-necessary explanation from molecular biology.[9]

So far, so good. The difficulties emerge when one probes how exactly the *relation of reduction* can be specified. As we have seen in the previous section, Rosenberg speaks about the "transformation" (2006, 53) or conversion (2006, 69f) of ultimate how-possible into proximate why-necessary explanations. However, as I have mentioned in Section 1.1.2, it remains obscure how this reduction process proceeds and what its result is. It appears to me that there are at least two different readings available:

First, Rosenberg states that during this reduction process "further historical facts – about genes and pathways – are added" and that "the links in the causal chain of natural selection are filled in" (2006, 46). In order to grasp what this amounts to, let us consider the example of the buckeye butterflies again. According to Rosenberg, the ultimate explanation of why buckeye butterflies developed eyespots on their wings is just a how-possible explanation that leaves several biologically pressing issues unexplained. For instance, it does not provide an answer to "which alternative adaptive strategies were available to various lineages of organisms,… how the feedback from adaptedness of functional traits – like the eyespot – to their greater subsequent representation in descendants was actually affected…[,] which feedback loops operate from fortuitous adaptedness of traits in one or more distantly past generations to improved adaptation in later generations, and how such feedback loops approach the biological fact to be explained as a locally constrained optimal design" (2006, 46f). In order to illuminate these issues we need to fill in the links in the causal chain with information about macromolecules. For example, we need to specify the strategies that were available for adaptation by identifying the genes that determine the characteristics of *Junonia coenia*'s ancestors and that provide the only stock of phenotypes

---

[7] More about Rosenberg's view on explanation can be found in Chapter IV.

[8] I am aware of the fact that Rosenberg does not use the term 'explanatory reduction'. However, he is concerned with reduction as a relation that exists between explanations. Hence, there is no reason why one should not classify this kind of reduction as explanatory reduction (together with individual reductive explanations, which will be discussed in Section 2).

[9] For the sake of simplicity I will sometimes also talk about higher-level explanations being reduced to lower-level explanations. However, I am aware of the fact that the distinction between higher-level and lower-level explanations does not completely coincide with the distinction between ultimate and proximate explanations.

on which selection can operate to move along pathways to alternative predation-avoiding outcomes. This talk of "adding information" and "filling in links of causal chains" suggests that something of the ultimate how-possible explanation is preserved during reduction, for instance, the general framework of the explanation. Otherwise it is unclear what this thing is to which information is added, or what is filled in.[10]

Second, other passages underpin a different reading. For example, when Rosenberg spells out the non-eliminative character of his reductionism he states that "functional biology's explanantia are always molecular biology's explananda" (2006, 54). Similarly, he characterizes reductionism as the thesis that "the complete or whole causal story is given at the level of macromolecules" (2006, 12). Passages like these suggest that the result of the transformation, respectively reduction process is an explanation, whose explanans *exclusively* refers to *molecules* and their interactions (i.e. to the genetic and biochemical pathway selection process) and does *not* invoke higher-level descriptions. According to this reading, nothing of the ultimate how-possible explanation is preserved during reduction. The only task that remains for functional biology would be the identification of the phenomena to be explained. What can be argued against this reading is that it is hard to imagine how molecular biology can develop explanations of higher-level phenomena completely on its own, that is, without relying on previously constructed higher-level explanations.

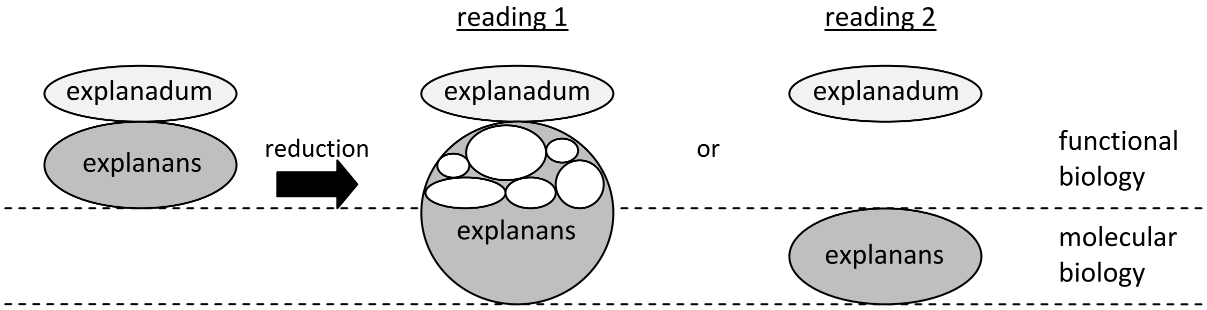

Figure III.1: Rosenberg's Notion of Explanatory Reduction

In sum, Rosenberg treats explanatory reduction as a relation between an ultimate and a molecular proximate explanation of the same phenomenon. The process of reduction he envisages is characterized as the transformation of the ultimate explanation into the proximate explanation in molecular terms, which involves elements of completion and correction. But whether some parts of the ultimate explanation are maintained during this transformation process and what, exactly, the result of this process is remains unclear.

---

[10] At this point one might doubt whether this "adding molecular details to an ultimate explanation" process really can yield a proximate explanation, as Rosenberg claims. I agree that this is a problematic step in Rosenberg's argumentation. However, one should note that Rosenberg states that "the reductionist's full explanation is still a historical explanation in which further historical facts – about genes and pathways – are added, and are connected by the same principles of natural selection that are invoked by the ultimate functional how-possible explanation" (2006, 46). But, in my view, this statement does not clarify much. Even worse, it completely blurs the difference between ultimate and proximate explanations.

1.3 Shortcomings of Rosenberg's Perspective

Rosenberg's defense of Darwinian reductionism is based on several difficult assumptions and involves many problematic argumentative steps (see e.g. Weber 2008; Love 2008b). However, in this section I will put these issues aside and focus on the difficulties his account of explanatory reduction encounters. One serious problem, of course, is the ambiguity of his account that I disclosed in the former section. It remains just too vague what the reduction of ultimate how-possible explanations to molecular proximate why-necessary explanations amounts to, what is preserved during this process, and what is the result of reduction. In what follows I will focus on three additional shortcomings of Rosenberg's perspective.

First, Rosenberg stresses that he wants to relinquish Nagel's model of theory reduction since it is "irrelevant" and does not "make contact with real issues in biology" (2006, 40). However, under closer inspection it turns out that Rosenberg does not move as far away from Nagel's model as this clear statement purports. Rosenberg's adherence to Nagel's model becomes evident in four different respects: First, Rosenberg considers not only the relation between explanations from different biological fields, but frequently invokes the relation between biological *theories*. This is illustrated, for instance, by the following citation: "Reductionism is the thesis that biological theories and the explanations that employ them do need to be grounded in molecular biology" (2006, 4). Second, just as the philosophers who tried to apply Nagel's model to biology (e.g. Kitcher 1984), Rosenberg focuses on the question whether functional biology's *generalizations* can be explained by molecular biology's generalizations in wide parts of his book (e.g. 2006, 32-39). Third, Rosenberg sticks to the D-N model of explanation and, in particular, to the view that *laws* are indispensable for the explanatory force of explanations (2006, Chapter 4).[11] Fourth, in accordance with Schaffner, Rosenberg focuses on *reduction in principle*. That is, he is neither interested in reconstructing real cases of epistemic reduction, nor in pursuing philosophy of biology in a bottom-up fashion (see Chapter I, Section 1.3) – even though Rosenberg explicitly claims that he wants to make contact with real issues in biology. Rather, it seems as if he conceives explanatory reduction in exactly the way he needs for his defense of Darwinian reductionism. In other words, he does not let biology speak to him, but takes out of biology what he needs (and thereby often misconstructs biology).

This adherence to Nagel's model makes Rosenberg's understanding of explanatory reduction vulnerable to all the criticism I raised in Sections 3.1 and 4 of Chapter II. All in all, what carries most weight is that Rosenberg does not pay much attention to what epistemic reduction in practice is, but rather creates an ideal image of explanatory reduction that fails to correspond with many aspects of actual biological practice. For instance, Rosenberg's account presupposes that for explanatory reduction to occur there must be a higher-level

---

[11] In Rosenberg's view there exists only one genuine biological law, namely the principle of natural selection, which is responsible for the explanatory force of all biological explanations (see 2006, Chapter 4).

and a lower-level (i.e. molecular) explanation for the same phenomenon available, so that the former can be converted into the latter. But in biological practice it very rarely is the case that two distinct explanations for the same phenomenon are developed, and that one of them is then transformed into the other. Rather, *single* reductive explanations are *directly* construed and this is the most common case of explanatory reduction.[12] This leads us to the second shortcoming.

Second, not only is Rosenberg's account of epistemic reduction inadequate with respect to biological practice, but so is the view of explanation it presumes. To begin with, Rosenberg's assumption that biology can be neatly separated into two fields that pursue two distinct explanatory tasks (i.e. functional biology individuates kinds functionally and develop ultimate explanations, whereas molecular biology individuates kinds primarily structurally and constructs proximate explanations) is a too simple and idealistic image of the structure of reasoning and knowledge in biology.[13] A considerable amount of biology is neither functional nor molecular and 'functional' means far more than 'evolutionary' in the biological sciences (which is why functional biology does not merely yield ultimate explanations). What is more, Rosenberg's claim that all ultimate explanations need to be converted into proximate explanations neglects the fact that both kinds of explanations are answers to different questions and, thus, conceptually independent and not competitors.

Finally, let me conclude with an objection that concerns Rosenberg's view of the relation between reduction and explanation. In a footnote he explicitly states:

> the debate [about reductionism] cannot be a dispute about 'explanation', for example a disagreement about pragmatic, erotetic, Protagorean versus nonerotetic accounts of explanation. For that is a general problem in the philosophy of science, not a problem about reductionism in the philosophy of biology. (2006, 41, footnote 6)

In my view it is a mistake to try to decouple the question of the correctness of reductionism from the question of the adequacy of explanations and Rosenberg's argumentation clearly shows why. But let us first track his argumentation for a moment. Rosenberg argues that Peter Railton's (1981) notion of an ideal explanatory text[14] enables us to avoid irrelevant debates about the nature of explanation since it directs our attention to the relevant issues (2006, 180). He claims that Railton's idea elucidates the following relevant difference between antireductionists and reductionists: the antireductionist holds that an ideal explanatory text need not advert to descriptions and generalizations about macromolecular processes and that a text adverting only to nonmolecular biological considerations could be ideal. The reductionist denies this thesis and insists that each ideal explanatory text will refer to macromolecular processes.

---

[12] This claim is supported by the various examples of paradigmatic and important cases of explanatory reduction I discuss in Chapter V.

[13] Love calls this an "artificial 'two-levellism'" (2008b, 3).

[14] Railton's notion of an ideal explanatory text is introduced and discussed in detail in Chapter IV.

I doubt that most antireductionists agree with Rosenberg on this point. Rather than claiming that there are also nonmolecular ideal explanatory texts, I think most antireductionists adopt a different argumentative strategy. They state that explanations are representations and, as such, always "partial" (e.g. Mitchell 2009, 23). Accordingly, explanations typically represent only parts of the ideal explanatory text. What distinguishes an antireductionist from a reductionist is that he claims that explanations, which refer only to nonmolecular parts of this text, will in many contexts be fully adequate. Hence, the antireductionist typically bases his argument on a pragmatic (i.e. erotetic) account of explanation (more on this in Chapter IV). This shows that the correctness of explanatory reductionism does in fact depend on what one regards as the conditions of the adequacy of an explanation and that one cannot decouple these two questions. It seems to me that Rosenberg's own defense of Darwinian reductionism even proves this thesis. Recall that his core assumption was that functional biology's ultimate how-possible explanations need to be reduced to molecular proximate why-necessary explanation because only the latter are maximally complete and, thus, adequate biological explanations. This thesis presupposes not only a nonerotetic account of explanation but also a very specific (and problematic) assumption about the constraints of the adequacy of explanations. Thus, Rosenberg himself links the question of reduction(ism) closely to the question of explanation.[15]

All in all, Rosenberg's account of explanatory reduction is not convincing. He fails to detach his account from the Nagelian framework. Furthermore, his perspective on explanatory reduction is shaped too much by the philosophical requirements his defense of reductionism carries with it. Thus, Rosenberg's account fails to capture what epistemic reduction and what explanation in contemporary biological practice really are.[16] In addition, he draws an artificial line between questions of reduction and questions of explanation that cannot be sustained – not even on the basis of his own argumentation.

2 Second Perspective: Individual Reductive Explanations

To regard explanatory reduction as a relation between two explanations is not the only available option. Those who adopt an alternative perspective take into account *individual* explanations and investigate the constraints of their reductive character. They are interested in questions like 'Under which conditions does a biological explanation succeed or fail to be reductive?' or 'What are the most important features of reductive explanations in the biological sciences?'. As I will elaborate in Chapter V, questions like these are based on the view that reduction is a relation between the two components of an explanation, namely between the explanandum (i.e. the phenomenon or behavior of a system to be explained)

---

[15] The close link between the question of reduction(ism) and the question of explanation does not imply that developing an account of explanatory reduction amounts to nothing more than developing an account of explanation. The contrary is the case, as I will elaborate in Chapter V.

[16] The empirical evidence for this claim can be found in Chapter V.

and the explanans (i.e. the explanatorily relevant factors).[17] In reductive explanations, the former is said to be reduced to the latter. What is characteristic for this perspective is that epistemic reduction is assumed to be intimately tied to specific types of explanation, namely to *part-whole explanations* (Kauffman 1970; Hüttemann/Love 2011) and to *mechanistic explanations*[18] (e.g. Wimsatt 1976a; Schaffner 2006, 385-390; Godfrey-Smith 2008, 56f; see also Chapter V, Section 5). In other words, part-whole and mechanistic explanations are treated as paradigmatic cases of reductive explanations.[19]

My goal in this section is to introduce and critically discuss the previous philosophical work on reductive explanations in the biological sciences. One important task is to identify fruitful insights that can be taken up in my own analysis of explanatory reduction. The central question to which I hope to find some viable answers is as follows: What is it that makes explanations in biology reductive? In other words, what are the features of reductive explanations that distinguish them from non-reductive ones? In order to find an answer to this question I first examine the work of the two precursors of this second perspective on explanatory reduction (Section 2.1), namely Kauffman's paper on part-whole explanations (1970) and Wimsatt's investigation of reductive explanations (1976a, 2007)[20]. Subsequently, I turn to the more recent work on reductive explanations. Although Kauffman and Wimsatt provided early valuable insights, Sarkar (1998, 2005) was the first philosopher of biology who gave a detailed analysis of the conditions that determine the reductive character of an explanation – at least with respect to explanations in genetics and molecular biology (Section 2.2). More recently, Hüttemann and Love (2011) took up this topic and examined different aspects of reductive part-whole reductive explanations in the biological sciences (Section 2.3). I conclude by summarizing the fruitful insights into the features of reductive explanations that have been offered so far (Section 2.4).

2.1 First Insights

The articles by Kauffman (1970) and Wimsatt (1976a) belong to those papers to which almost everybody who writes about part-whole, mechanistic, or reductive explanations refers, but only few discuss in detail. This is reason enough to take a look at some of the details.

---

[17] More precisely, it is the *description* or representation of the phenomenon that is being reduced to the *description* or representation of the explanatorily relevant factors.

[18] For instance Wimsatt writes: "At least in biology, most scientists see their work as explaining types of phenomena by discovering mechanisms… and *this* is seen by them as reduction" (1976a, 671).

[19] In which respects reductive explanations can deviate from part-whole explanations and from mechanistic explanations will be exposed in Chapter V.

[20] Wimsatt's paper "Reductive Explanation: A Functional Account" (1976a) is reprinted as a slightly modified version in his newest book "Re-engineering Philosophy for Limited Beings" (2007). In what follows I cite only the original paper.

2.1.1 Kauffmann's Analysis of Part-whole Explanations

The primary target of Kauffman's analysis is not reductive explanation, but part-whole explanation (or, as he calls it, "parts explanation"). However, since part-whole explanations are paradigmatic cases of reductive explanations, it can be expected that an analysis of the former also provides insights into the latter. Kauffman offers the following characterization of part-whole explanations. In his view, the explanandum of a part-whole explanation is a specific behavior of a system ("what the system is doing" 1970, 259) and the part-whole explanation is an explanation of *how* the parts and processes of this system articulate together to cause the system's behavior ("the interworking of its parts" 1970, 258).

Large parts of Kauffman's paper are dedicated to seeking an understanding of the process by which biological part-whole explanations are generated.[21] According to his analysis, the first step of this process is to single out the behavior to be explained from the various behaviors a system displays. After that, the system is decomposed into those parts and processes that fit together to yield the behavior in question. Kauffman spends much time on pointing out the various ways plurality comes into play in this process of developing a part-whole explanation. He argues that the decomposition of a system does not only depend on the characterization of the behavior of the system to be explained. Moreover, if "diverse sets of sufficient conditions" (1970, 259) for what an adequate description is, are applied, a system can also be decomposed differently with respect to the same behavior. Kauffman states that this plurality of possible decompositions of a system gives rise to a plurality of part-whole explanations, which cannot be brought under "some overarching, ultimate view of what the organism is 'really' doing" (1970, 272).

Two other ideas of Kauffmann's are worth mentioning here. The first is his claim that the development of a part-whole explanation does not only involve the identification of the parts of a system. In addition, he claims that one needs to identify the relevant "causal consequences" (1970, 260) of the parts. Consequently, part-whole explanations provide a view of "what it is that the parts… [are] doing from among the indefinitely many possible things each part might… be doing" (1970, 265). This claim is similar to the thesis of the New Mechanists that mechanistic explanations refer to both, to entities (i.e. the spatial parts of a system) as well as to what the entities are doing (i.e. their interactions, activities, operations, respectively the temporal parts of the system; see also Chapter V, Section 1.2.2).

The second idea concerns the interdependency between the two tasks of decomposing the system and of characterizing the behavior of the system to be explained. According to Kauffman, on the one hand, the decomposition of a system is guided by what is regarded as the behavior of the system to be explained (1970, 259f). For instance, you will decompose a cell differently if the relevant behavior is the synthesis of proteins or if it is the division of the cell. On the other hand, new information about the parts of a system may

---

[21] The great interest in the *discovery* of explanations is representative for the New Mechanistic Philosophy, too (see, in particular, Craver/Darden 2001, 2005; Craver 2002b; Bechtel 2006).

also call for a revision of the characterization of the system's behavior (1970, 269f). For instance, Harvey's discovery that the blood circulates gave rise to a new view of what the heart does. The changed characterization of the system's behavior can in turn lead to a revised identification of parts. This interplay between the description of the phenomenon to be explained and the characterization of the parts of a system (respectively the parts of a mechanism) is also something the New Mechanists emphasize (e.g. Craver/Darden 2001, 119-123; Bechtel 2006, 28-33).[22]

All in all, Kauffman's paper presents an important analysis of the nature of part-whole explanations and of the process of developing them via decomposition (see Chapter II, Section 3.2.2). But besides the fact that his paper sheds light on a paradigmatic case of reductive explanations, it contains no answer to my central question of what it is that makes part-whole explanations *reductive*. Hence, Kauffman's ideas do not advance my project of specifying the reductive character of biological explanations.

2.1.2 Wimsatt on Reductive Explanations

Let us turn to the second paper which is assigned a classical status in the debate. At first sight, Wimsatt's paper seems to be more fruitful for my purpose than Kauffman's since he explicitly addresses the topic of reductive explanations – this is at least what the title "Reductive Explanation: A Functional Account" suggests. However, if one takes the trouble to work through the abundance of ideas and arguments in Wimsatt's paper one will notice that he primarily provides a criticism of Nagel's "standard model" (1976a, 681) of reduction, rather than an analysis of the features of reductive explanations. I refrain from reviewing Wimsatt's arguments against Nagel's model here. The most important ones have found their way into my critical discussion of the Nagelian account of theory reduction in Section 4 of Chapter II. What interests me at this point is whether Wimsatt's paper yields any insights into the features of reductive explanations in the biological sciences.

Due to the fact that Wimsatt's primary focus is to criticize Nagelian accounts of reduction, only a few comments on what makes explanations reductive can be found throughout his paper. According to Wimsatt, in reductive explanations "the properties of higher level entities are… explained in terms of the properties and interrelations of *lower level* entities" (1976a, 680; my emphasis). Elsewhere, he explains that the mechanisms and causal factors cited in reductive explanations "are at a *lower level* of organization than that of the phenomenon being explained" (1976a, 686; my emphasis). For this reason, Wimsatt calls reductive explanations also "micro-level… explanation[s]" (1976a, 689) and contrasts them with two other types of explanations: on the one hand, with explanations that show how the phenomenon is a product of causal interactions at its own level and, on the other hand, with explanations that show how the phenomenon is a product of causal interactions

---

[22] It is interesting and in a way surprising that many of the ideas of the New Mechanists, which at first sight appear to be brand-new, already can be found in Kauffman's paper.

at higher levels of organization. The latter are characterized as functional explanations (1976a, 689).

To conclude, Wimsatt provides an important first insight into what it is that marks the reductivity of explanations. His view can be summed up by the following characterization of reductive explanations:

> Explanations are reductive iff the factors cited in the explanans are located on a *lower level* of organization than the phenomenon to be explained.

Let us now turn to the work of Sarkar and explore a more in-depth analysis of the conditions that determine the reductive character of an explanation.

2.2 Sarkar's Analysis of Reduction in Genetics

In his book "Genetics and Reductionism" (1998), and in his collection of papers called "Molecular Models of Life" (2005), Sarkar explicitly addresses the question what criteria an explanation must satisfy to constitute a reduction. The core of his analysis are three criteria of reductivity, referred to as "*fundamentalism*", "*abstract hierarchy*", and "*spatial hierarchy*" (1998, 43f). According to these criteria Sarkar distinguishes three main types of (explanatory) reduction that frequently arise in genetics and molecular biology. But before I discuss the core of Sarkar's account in detail (Section 2.2.3), let us have a look at the general framework of his analysis. Two issues are of particular importance: on the one hand, Sarkar's characterization of his approach as substantive (Section 2.2.1) and, on the other hand, his claim that an account of explanatory reduction must be independent from an account of explanation (Section 2.2.2).

2.2.1 Formal vs. Substantive Analyses of Epistemic Reduction

Sarkar characterizes his analysis of reduction in genetics as non-formal and "substantive" (1998, 18) in order to clearly demarcate it from previous accounts of reduction (like Nagel's) that have focused on formal issues. Highlighting this difference is so important to Sarkar that it determines the structure of his book (1998).

However, the distinction between formal and non-formal accounts of epistemic reduction is not new. For instance, Wimsatt characterizes his approach as "functional" (1976a, 675, 700) and contrasts it strongly with Nagel's formal model of reduction. Wimsatt argues that it is inadequate to reconstruct epistemic reduction exclusively by considering its logical structure. Instead, epistemic reduction should be analyzed with respect to its functioning in promoting the aims of science (most notably explanation), which is why Wimsatt refers to his account as 'functional' (1976a, 673-675, 700). One might go even further and try to trace back the difference between formal and non-formal analyses of epistemic reduction to Nagel's distinction between formal and non-formal conditions for reduction (1961, 345-380).[23]

In accordance with Wimsatt, Sarkar emphasizes the non-formal character of his approach. But he spells out the meaning of 'non-formal' differently. Sarkar provides the following explication:

> Syntactic issues and those semantic issues that have been framed as questions of form [e.g. whether bridge statements are identity claims] will be called 'formal' here. […] Other (scientifically or philosophically) relevant issues, which will generally concern the interpretation of scientific arguments, especially what they imply and what they assume about the world, will be called 'substantive'. (1998, 19)

According to this quote, models of epistemic reduction like Nagel's are formal since they characterize reduction as a purely logical or semantical-analytical relation between (sets of) statements or terms. In contrast, a substantive account of reduction ignores the logical features of reductive statements and focuses on the question what reductive statements presume about the entities and relations that exist *in the world.*[24] Unfortunately, this is all what Sarkar says about the substantive character of his approach. He leaves unanswered the question of what exactly it means to analyze reductive explanations with respect to the substantive claims they make about the world.

One might suggest that to pursue a substantive analysis means to shift the focus from epistemic reduction to ontological reduction (see Chapter II, Section 2). I think it is important to note that this is *not* what it means. Even if Sarkar focuses on the question of what reductive explanations assume about the world, he is still concerned with reduction as a relation between *epistemic* units (in his case, the parts of an explanation). Rather, what makes his analysis substantive (or in my words "ontic"; see Chapter V, Section 6) is his claim that the reductivity of an explanation is determined by the *assumptions about the world* that are made in explanations (e.g. assumptions about causal relations, levels of organization, system-environment distinctions, etc.). This thesis can be illustrated by Sarkar's three criteria of reductivity (which will be discussed in detail in Section 2.2.3). These criteria (i.e. fundamentality, abstract hierarchy, and spatial hierarchy) can be formulated as relational

---

[23] However, such an attempt is problematic in two respects: on the one hand, Nagel's notion of non-formal conditions for reduction does not concur with the non-formal character of Sarkar's or Wimsatt's account. According to Nagel, the additional non-formal conditions for reduction serve to distinguish trivial from noteworthy scientific achievements since they require that theoretical assumptions must be supported by empirical evidence and must exhibit some degree of predictive power (1961, 358-361). This is not the same as what Sarkar and Wimsatt have in mind when they speak about the non-formal character of their account of reduction. On the other hand, Nagel's non-formal conditions do not suffice on their own to distinguish cases of epistemic reductions from non-reductions. This is due to the fact that they hold for all theoretical assumptions in science. Hence, there is a good reason why Sarkar and Wimsatt do not adopt Nagel's notion of non-formality: his non-formal conditions for reduction fail to provide a basis for an adequate non-formal analysis of epistemic reduction.

[24] This is in line with Sarkar's other remarks on the substantive character of his account. For instance, he refers to his criteria of reductivity as 'substantive' because "they are about what assumptions are made during a (putative) reductionist explanation, rather than about the form that such an explanation may take" (1998, 43).

predicates: ‘is fundamental to’, ‘is located on a lower level of an abstract hierarchy than’, and ‘is located on a lower spatial level than’[25]. What makes these criteria substantive is the fact that whether they are fulfilled or not depends on the causal and constitutional relations that exist in the world (and, of course, on how these relations are represented in the explanation). Consider the example of light reaction of photosynthesis. The behavior of a chloroplast to transform light energy into chemical energy is reductively explained by describing how photons are absorbed by photosystems, how electrons are excited and flow down an electron transport chain, how this creates a proton gradient across the chloroplast membrane, which is then used to synthesize ATP. According to Sarkar’s analysis this explanation is reductive since it fulfills the criterion “spatial hierarchy”. This means that the system, whose behavior is to be explained, is represented as being located on a *lower spatial level* than the entities and processes referred to in the explanans. Thus, the existence of a part-whole relation between *entities in the world* – that is, between the entity described in the explanandum (namely the chloroplast, respectively the process of light reaction) and the entities described in the explanans (namely the photosystems, electrons, chloroplast membrane, ATP synthase, etc. as well as what these objects are doing, e.g. absorbing, being excited, transforming, etc.) – is assumed to determine the reductive character of the explanation. This is what makes Sarkar’s account substantive.

2.2.2 Analyzing Reductive Explanations without Specifying Explanation

The second notable feature of Sarkar’s analysis is that he makes great efforts to keep his account of explanatory reduction *independent* from any particular model of explanation (1992, 178f; 1998, 9f, 39-43). Sarkar emphasizes that he aims at identifying

> “*additional* criteria that the explanation must satisfy in order to be a reduction” and that he wants to “make sure that the criteria for reduction remain general enough that they are not likely to fall afoul… of usual explications of explanation” (1998, 9).

The attempt to separate the question of reduction from the question of explanation is comprehensible. If it succeeds it will save one from getting entangled in the controversial debate about what an explanation is. I strongly endorse Sarkar’s call for additional criteria of reductivity, which distinguish reductive from non-reductive explanations without falling together with those criteria that distinguish explanations from non-explanations (e.g. pure descriptions). However, I am also less optimistic than Sarkar that the neutrality or compatibility of an account of reduction with regard to competing accounts of explanation that he demands can be satisfied. To begin with, Sarkar’s analysis itself is not completely neutral on this point. For instance, he admits that an analysis of reductive explanations cannot be pursued without *any* specifications of what explanations are. As a result, he presents four “basic assumptions” (1998, 41) about explanations he presumes.

---

[25] Objects on different spatial levels are related to each other via part-whole relations or relations of spatial containment, e.g. parts of a system are located on a lower spatial level than the system (Sarkar 1998, 55).

(1) Explanations are *representations* of systems. The same system can be represented in different ways (depending on the context of investigation).
(2) The phenomenon to be explained is some *feature* of the system (which is typically called 'behavior of the system') as represented, for example a law it obeys or an event in which it participates.
(3) An explanation involves a "process of scientific reasoning or explanation… called a *derivation*" (1998, 42). The degree of precision and the type of mathematical rigor of the derivation depends on the scientific context.
(4) Any explanation entails a set of *explanatory factors*, which are presumed to be the relevant ones. These factors can, but need not be referred to in a general law. The relevance of explanatory factors is determined by the context.

I agree with Sarkar that most of these assumptions are quite neutral, that is, they are compatible with different models of explanation. Yet the third assumption appears to be an *ad-hoc* assumption, which is too vague to be convincing. In order to retain the neutrality of his account Sarkar emphasizes that the term 'derivation' should not be equated with the logician's notion of derivation (since this would commit him to a D-N model of explanation). However, it remains obscure what this wider notion of derivation is and whether and how it can be distinguished from the general notion of reasoning (see also Chapter II, Section 4.2).

The fourth assumption reveals that Sarkar does not restrict the notion of an explanation to explanations that cite laws in the explanans. He allows that the set of explanatorily relevant factors may be described by generalizations that are not laws or that they may be only particular factors (tokens). However, Sarkar's other remarks convey a different picture. When he specifies his first criterion "fundamentality" he states that in reductive explanations the assumption is made that "the feature to be explained is a result only of the *rules* operative in that realm [i.e. in the fundamental realm]" (1998, 43; my emphasis; see also 1998, 46). This quote implies that the explanans in reductive explanations must *always* contain generalizations that describe the rules or regularities on the fundamental level. Although this does not commit Sarkar to the D-N model of explanation, it excludes explanations that refer only to particular objects and processes. However, one could positively interpret the quoted passage as merely claiming that many (and not all) reductive explanations cite "rules". And since most biological explanations in fact involve generalizations, this objection proves not to be very compelling.

The criticism has so far challenged the neutrality of Sarkar's basic assumptions about explanations. Besides this another more general and, I think, more serious concern can be raised with regard to Sarkar's requirement that an analysis of explanatory reduction should be compatible with any model of explanation. This concern arises from Sarkar's arguments against the adequacy of a formal account of reduction (see Section 2.2.1). As it seems to me, there exists an inevitable tension between, on the one hand, Sarkar's demand for the neutrality of an account of reduction with respect to different models of explanation and, on the other hand, his strict rejection of formal accounts of reduction. In the context of reduction Sarkar explicitly dismisses formal models of reduction that reconstruct reduction with respect to its logical features. Against this background, the possibility that Sarkar at the same time adopts the D-N model of explanation seems weird because the D-N model traces back the explanatory force of an explanation (also) to its formal features (i.e. to the logical relation between the statements of an explanation). Thus, the combination of Sarkar's *non-formal* account of reduction with the *formal* D-N model of explanation seems to be untenable. At least, a substantive account of reduction goes much more smoothly along with a "substantive" account of explanation, for instance, with an "ontic conception" of explanation (Salmon 1993, 80) that attributes the explanatory force to the embedding of a phenomenon into the causal structure of the world. Hence, Sarkar would have done well to give up his demand for neutrality and to accompany other opponents of Nagel's model, like Hull (1974) and Wimsatt (1976a), who endorse a non-formal account of explanation.

2.2.3 Criteria of Reductivity

Let us turn to what I take to be the core of Sarkar's account of explanatory reduction, namely his answer to my central question what is responsible for the reductive character of an explanation. Sarkar develops three substantive criteria of reductivity, by means of which one can assess the reductive character of an explanation:

> (i) *Fundamentalism:* the explanation of a feature of a system invokes factors from a different realm (from that of the system, as represented) and the feature to be explained is a result only of the rules operative in that realm.
>
> (ii) *Abstract hierarchy:* the representation of the system has an explicit hierarchical organization, with the hierarchy constructed according to some independent criterion (that is, independent of the particular putative explanation), and the explanatory factors refer only to properties of entities at lower levels of the hierarchy.
>
> (iii) *Spatial hierarchy:* the hierarchical structure referred to in (ii) is a hierarchy in physical space; that is, entities at lower levels of the hierarchy are spatial parts of entities at higher levels of the hierarchy. The independent criterion invoked in (ii) now becomes spatial containment. (1998, 43f)

From Sarkar's remarks it can be inferred that the first criterion, "fundamentalism", is sufficient as well as necessary for an explanation to be (weakly) reductive. In other words, Sarkar claims that any reductive explanation in genetics and in molecular biology satisfies at least this criterion. The other two criteria "abstract hierarchy" and "spatial hierarchy" can be met in addition, whereas the fulfillment of the third criterion always implies that the second criterion is also satisfied because a spatial hierarchy is a special case of an abstract hierarchy. In the following I will explicate and discuss these three criteria of reductivity in more detail.

The first criterion, "fundamentalism", captures Sarkar's intuition that reduction has something to do with identifying two realms, treating one of these realms as epistemologically prior (i.e. as fundamental), and tracing back a part of the non-fundamental realm to a part of the fundamental realm. In the case of reductive explanations the description of the phenomenon to be explained is reduced to the explanatory factors cited in the explanans in the sense that the phenomenon belongs to the *fundamental realm*, whereas the explanatory factors refer to entities of the non-fundamental realm, and the former is traced back to (i.e. explained by) the latter. Sarkar calls a reductive explanation that satisfy only this first criterion a "weak reduction" (1998, 44).

The criterion "abstract hierarchy" adds another requirement to the first criterion. In reductive explanations that satisfy this second criterion the two different realms (to which the explanandum and the explanans refer) do not stand side by side, but are ranked hierarchically. "Abstract hierarchy" can be interpreted as specifying the sense in which one realm is more fundamental than the other. If an explanation represents an abstract hierarchy[26] (like the allele-genotype-phenotype hierarchy) one realm can be said to be more fundamental than the other in the sense that it constitutes a *lower level* than the level the other realm constitutes. For instance, genotypes can be said to belong to a lower level of an abstract hierarchy than phenotypes. What distinguishes reductive explanations that satisfy only this criterion from reductive explanations that fulfill also the third, "spatial hierarchy" criterion is that the hierarchy they represent is not a spatial hierarchy, but only an "abstract hierarchy" (1998, 53-55).[27] A typical example of a reductive explanation that satisfies the first and second criterion is the explanation of a phenotypic trait of an organism (e.g. the green seed color of peas, the sickle cell trait or Huntington's disease of humans, etc.) by reference to certain genes and a plausible model of their transmission. According to Sarkar's analysis, explanations like these are reductive because genes, respectively alleles (explanans), are located on a lower level of an abstract hierarchy than phenotypes (explanandum). Sarkar names this kind of reductive explanation "abstract hierarchical reduction" (1998, 44). An important result of his analysis is that in classical genetics these abstract hierarchical reductions (which he also calls "genetic reductions"; 1998, 103) are the most common kind of reductive explanations.[28]

Reductive explanations that satisfy all three criteria are by far most common (in genetics and molecular biology). They are characterized by a hierarchy, which is spelled out spatially (thus, the third criterion is named "spatial hierarchy"[29]). In reductive explanations

---

[26] Sarkar argues that abstract hierarchies need to be "constructed according to some independent criterion" (1998, 43). This condition serves to preclude rare cases, in which the hierarchy is postulated only for the sake of the explanation.

[27] According to Sarkar it is irrelevant that at least a part of the allele-genotype-phenotype hierarchy could in principle be spelled out spatially (since genomes spatially consist of linkage groups, which in turn consist of loci that are occupied by alleles). What is important is not whether the hierarchy *might* to some extent be spelled out spatially, but that the hierarchy is *in fact* represented as an abstract and not as a spatial hierarchy in reductive explanations in genetics (1998, 128).

[28] It is important to note that this characterization of the reductivity of genetic explanation does not imply the admittedly controversial thesis that reductions of this kind will always be successful. Particularly, in the case of mental disease traits or other complex behavioral traits the development of a genetic explanation is highly problematic if not impossible. For a detailed discussion of these problems compare Sarkar 1998, 124-127, 131f.

[29] In a later paper Sarkar refers to this criterion also as the "criterion of compositionality" (2008, 429).

of this kind the entities referred to in the explanans belong to the fundamental realm in the sense that they are located on a *lower spatial level* than the phenomenon represented by the explanandum. To put it another way, the entities referred to with the explanatory factors are spatial parts of the system whose behavior is to be explained. Reductive explanations of this third kind are also called part-whole explanations. A typical example is the explanation of DNA replication in cellular organisms (which can be conceptualized as a behavior of cells) by reference to the separation of the DNA strands, the synthesis of RNA primers, the moving of DNA polymerase along the DNA strand, the continuously adding of nucleotides, the removal of the RNA primers of the lagging strand, etc. This explanation is reductive since it satisfies Sarkar's "spatial hierarchy" criterion (together with the other two). It explains a certain behavior of a cell (i.e. DNA replication) by describing only spatial parts of the cell (i.e. DNA strands, DNA polymerase, RNA primer, etc), the way they are organized, and how they interact with each other. Reductive explanations like this are called "strong reduction[s]" (1998, 44). Sarkar stresses that, contrary to in classical genetics, the reductive explanations that can be found in molecular biology belong to this kind of strong reductions (also referred to as "physical reductions"; 1998, 136).

In sum, Sarkar provides the following characterization of reductive explanations (let P be the phenomenon or behavior of a system to be explained (represented in the explanandum) and let E be the entities referred to with the explanatorily relevant factors (represented in the explanans):

Explanations are reductive iff
(1) P and E belong to different realms, whereas the realm of E is assigned a fundamental status (*criterion of fundamentalism*)
or
(2) criterion (1) is fulfilled, the explanation includes the representation of an abstract hierarchy, and E is located on a lower level than P (*criterion of abstract hierarchy*)
or
(3) criteria (1) and (2) are fulfilled and the hierarchy referred to in (2) is spelled out spatially (*criterion of spatial hierarchy*).

2.2.4 Shortcomings of Sarkar's Account

As we have seen, Sarkar provides an elaborate substantial analysis of the reductive character of explanations in genetics and molecular biology. In comparison with Wimsatt's few remarks on the nature of reductive explanations, Sarkar's analysis yields some important additional insights. Whereas Wimsatt recognizes only strong reductions (i.e. reductive explanations that meet the criterion "spatial hierarchy"), Sarkar identifies two other kinds of reductive explanations in biology. This seems to be a valuable step towards capturing the diversity of reductive explanations in biology. However, Sarkar's account has its shortcomings, too. In what follows, I will highlight three of them.

First, Sarkar's analysis is restricted in scope since it takes into account only reductive explanations in genetics and molecular biology. Although this allows an in-depth investigation of the explanatory practice of these two disciplines, it has the apparent drawback that it sheds light only on a small range of reductive explanations in biology. Accordingly, Sarkar's account might fail to capture the diversity of reductive explanation present in contemporary biology (see Chapter I, Section 4).

Second, due to his focus on molecular biology Sarkar offers a too one-sided interpretation of his "spatial hierarchy" criterion. In wide swathes of his work he considers only part-whole explanations in which the identified parts are located exclusively on the molecular level.[30] In doing so, he ignores part-whole explanations, in which the parts are located on a lower level than the phenomenon to be explained, but not on the lowest level of molecules (which I call *the* one fundamental level). As I will elaborate in Chapter V, Section 2.3.1, such a restriction of the notion of reductive explanations to the special class of molecular explanations (i.e. to what I call fundamental-level reductions) is inadequate and has proved to be misleading in several contexts.

Third, Sarkar's "fundamentalism" criterion is not convincing – at least not if it is conceived as a discrete sufficient criterion that distinguishes a significant class of reductive explanations. On the one hand, the class of weak reductions appears to be almost empty. Sarkar himself only marginally discusses examples (like genetic explanations, in which no particular structure is attributed to the genotype at all; see 1998, 71; and reductive explanations in ecosystem ecology; 2005, 111f). All in all, weak reductions do not seem to constitute an important research aim in biology. On the other hand and more importantly, "fundamentalism" is not a sufficient criterion for distinguishing reductive from non-reductive explanations. The reason is that any explanation (i.e. also the non-reductive ones) involves decisions about explanatory primacy (i.e. about explanatory relevance).[31] Since explanations are representations and as such only partial, any explanation involves distinguishing some factors as explanatorily prior or as explanatorily relevant and dismissing other factors as irrelevant. And exactly this preference of some factors over others is what Sarkar means (unlike I) by distinguishing something as fundamental or as non-fundamental.[32] Thus, it seems as if Sarkar's account results in the implausible view that almost all explanations are reductive since almost all explanations involve distinguishing some factors as explanatorily

---

[30] A notable exception constitutes Sarkar's philosophical work on ecology (e.g. Sarkar 2005, 111f; 2009).

[31] In a recent paper Sarkar seems to recognize this drawback (at least with respect to research strategies): "All potentially non-reductionist research strategies discussed satisfy the criterion of *epistemic primacy* required for weak reductions." (2008, 433)

[32] One might object that characterizing a factor or a kind of factors (or a realm) as fundamental involves more than just characterizing it as explanatorily relevant or prior. This is true for our intuitive understanding of the term 'fundamental'. However, the way Sarkar spells out his "fundamentalism" criterion lacks this supplement. He leaves it almost completely unspecified what it means for one realm to be more fundamental than another. Different qualifications of 'fundamental' come into play only in his two other criteria. In this context he specifies that 'more fundamental' means 'being located on a lower level of an abstract/spatial hierarchy'.

prior or as fundamental and ignoring other factors as irrelevant. In contrast, I think that fundamentalism (or explanatory primacy) is not sufficient to demarcate reductive from non-reductive explanations.

2.3 Hüttemann's and Love's Three Aspects of Reductive Explanation

Given the popularity of the topic of reduction(ism) it is surprising that only few philosophers have devoted themselves to the task of scrutinizing the nature of reductive explanations in biology. A recent notable exception is the paper of Hüttemann and Love on "Aspects of Reductive Explanation in Biological Science: Intrinsicality, Fundamentality, and Temporality" (2011).[33] The focus of this paper is on disclosing the different respects in which a biological explanation can succeed or fail to be reductive. The case study they consider in detail is "the protein folding problem" (2011, 534) in molecular biology.

In this section I present what I regard as the major results of Hüttemann's and Love's analysis and discuss the difficulties it encounters. I start with some brief remarks on their methodological framework (Section 2.3.1). Then I turn to the key part of their analysis. I explicate what they mean by the three aspects of reductive explanation (i.e. intrinsicality, fundamentality, and temporality) and how these aspects arise from distinguishing two core facets of reductive reasoning, namely composition and causation (Section 2.3.2). In the subsequent section I point out Hüttemann's and Love's answer to my central question: What constitutes the reductive character of biological explanations (Section 2.3.3)? I conclude by revealing some objections their account faces (Section 2.3.4).

2.3.1 Methodological Framework

One reason why the analysis of Hüttemann and Love is of great value to my project is that they aim at a target that is similar to mine. Furthermore, they take for granted similar methodological assumptions as I do (see Chapter I). Their self-proclaimed goal is

> to understand and explicate the diversity of reductionist reasoning practices in biology ('application'), with special attention to their utilization in ongoing research ('relevance') (2011, 521).

Hence, Hüttemann and Love share my goal of understanding what reductive explanations in current biological practice in fact are, that is, what the most significant characteristics of reductive explanations in biology are and according to which criteria or constraints they are distinguished from non-reductive explanations. Their demand for "application" corresponds to my criterion of descriptive adequacy (Chapter I, Section 1) and their requirement of "relevance" is in line with what I characterized as the potential usefulness of philosophy to biological practice (Chapter I, Section 6).

---

[33] A modified version of this paper, which is focused more on the differences between part-whole explanations in biology and in physics, was published in the same year (Love/Hüttemann 2011).

However, the goal Hüttemann and Love seek to achieve deviates from my own in a significant respect. They stress that they do not want to "produce an overarching theory of reductive explanation as a competitor to GRR models [i.e. Nagel-Schaffner models of theory reduction] or other accounts" (2011, 521). Rather, they aim at disclosing different "aspects" of reductive explanations, thereby illuminating the heterogeneous nature (or diversity) of reductive explanations in biological science (2011, 524). Compared to this, my own project seems to be more ambitious. I seek to develop an account of epistemic reduction that constitutes an *alternative* to earlier models, including Nagelian models of theory reduction. Hence, I do not shrink from calling my account a "theory" of reductive explanations.[34]

2.3.2 Intrinsicality, Fundamentality, and Temporality

At the outset of their paper Hüttemann and Love identify two "major facets" (2011, 522) of reductive explanation, namely *composition* and *causation*. They point out that reductive explanations in biology typically involve compositional claims as well as causal claims. For instance, the explanation of protein synthesis contains, on the one hand, the claim that cells (i.e. the system whose behavior is to be explained) are composed of or constituted by certain cell organelles (e.g. ribosomes, the nucleus, etc.) and certain macromolecules (e.g. DNA strands, amino acids, m- and t-RNAs, etc.). On the other hand, it comprises claims about causal processes like the thesis that the synthesis of proteins is brought about by a sequence of interactions between the various components of the cell (e.g. the binding of the ribosome subunits to the m-RNA strand, or the termination of the polypeptide when a stop codon is reached, etc.). As Hüttemann and Love persuasively show for the case of protein folding, from a philosophical perspective it is often profitable to tell apart questions about causation from questions about composition. For example, the question of whether native protein conformation can be inferred from the amino acid sequence of the polypeptide (what is named the "predictive construal" of the linear sequence hypothesis; 2011, 542) needs to be distinguished from the question of whether the causal process of how the folding occurs can be explained solely in terms of the properties of amino acids (i.e. the "folding construal" of the linear sequence hypothesis; 2011, 542). Hüttemann and Love stress that a positive answer to the former does not imply an affirmative answer to the latter.

The distinction between these two facets of reductive explanations (i.e. composition and causation) provides the basis on which Hüttemann and Love determine three "key aspects" (2011, 523) of reductive explanations in biology. Before I go into the details of each of these three aspects, let me add a general remark. What is crucial for Hüttemann's and Love's analysis is that they treat reductive explanations as equal to "compositional part-

---

[34] In fact, I do not see a good reason why Hüttemann and Love should not do that, too. Taking into account the diversity of reductive reasoning and seeking an account of reductive explanation that is sensitive to actual biological practice contradicts neither calling this account a "theory" of reductive explanations nor regarding it as an alternative to other accounts of epistemic reduction.

whole reductions" (2011, 527),[35] that is, to *part-whole explanations* in which the behavior or a property of a whole is explained in terms of the properties of its parts.[36] The first two of Hüttemann's and Love's aspects, namely intrinsicality and fundamentality, are supposed to capture two distinct respects in which these part-whole explanations typically are reductive:

> First, they appeal solely to *intrinsic* features of the compound system in question [i.e. to the parts of the system]…. Secondly, they appeal to a more *fundamental* realm or lower level features (the parts), or a restricted set of properties within this realm, as compared with the whole (the non-fundamental realm). (2011, 527f)

By contrast, the third aspect, namely temporality, is not said to be related to the part-whole character of reductive explanations since it is associated with causation and not with composition (2011, 523f). Let us now have a look at how these three aspects are specified.

Intrinsicality

According to the analysis of Hüttemann and Love part-whole explanations typically are reductive since they solely refer to *intrinsic* features of the system (whose behavior is to be explained), namely to the parts of the system and to their properties. In other words, reductive explanations appeal only to factors that are "*contained within*" or "*internal to*" (2011, 528) the outer boundary of the system in question. Hüttemann and Love claim that this focus on the intrinsic features of a system presupposes that, at first, the system of interest is identified, which includes its demarcation from the environment or surrounding context. They emphasize that "natural phenomena do not come with labels attached to indicate their boundaries" (2011, 528) and that how the boundary between system and environment is drawn (i.e. what is regarded as intrinsic and extrinsic to the system), depends on the explanatory goals of researchers. Thus, they conclude that "[i]ntrinsicality has an epistemological (or pragmatic) aspect" (2011, 523).

Fundamentality

The second aspect of reductive explanations Hüttemann and Love identify is fundamentality. According to this, part-whole explanations typically are reductive in the sense that they appeal only to *more fundamental* features or properties (I would say factors). The way this is usually spelled out in the context of part-whole explanations is that the parts of a system are more fundamental than the whole system in the sense that they are located on a *lower level* of a spatial hierarchy. Recall Sarkar's criteria of reductivity (Section 2.2.3). This is exactly what Sarkar aims to capture by his "spatial hierarchy" criterion.

---

[35] However, in Section 2.3.3 I reveal that it remains unclear whether Hüttemann and Love identify reductive explanations with part-whole explanations or whether they assume that there exist non-reductive part-whole explanations, too, as some of their remarks suggest.

[36] Although I agree with them that part-whole explanations are paradigmatic cases of reductive explanations, in my view it is also important to recognize other features of reductive explanations than their part-whole character as well as other kinds of reductive explanations (see Chapter V).

However, the way in which Hüttemann and Love explain what they mean by 'fundamentality' displays that they admit different respects in which the explanatory factors can be said to be more fundamental than the phenomenon to be explained. In the above quotation they state that "they [i.e. the explananda of reductive explanations] appeal to a more fundamental realm or lower level features…, or a restricted set of properties within this realm" (2011, 528). This suggests that Hüttemann and Love accept three different readings of 'more fundamental': first, Sarkar's criterion "fundamentalism", according to which 'being more fundamental' means 'belonging to a more fundamental realm' (without specifying what makes this realm more fundamental); second, Sarkar's criterion "spatial hierarchy", according to which 'being more fundamental' means 'being located on a lower spatial level'. Third, they introduce a "qualified sense… of fundamentality" (2011, 524), according to which 'being more fundamental' means 'belonging to a restricted set of factors', for instance, to the set of biochemical properties. One difficulty of Hüttemann's and Love's paper is that they are not always clear about the notion of fundamentality they actually use. Sometimes they indicate that they refer to the third notion by using the phrase "qualified fundamental level" (2011, 529) or "qualified sense of fundamentality" (2011, 538). However, in the majority of cases they only talk about the fundamental realm, fundamental levels, or fundamentality in general.

This is problematic in at least two respects: first, although Hüttemann and Love sometimes use the term 'fundamental realm' without specifying this fundamentality in a spatial-hierarchical manner, it remains obscure whether they really adopt Sarkar's criterion "fundamentalism". The reason is that they consider only part-whole reductions and, as such, only cases in which the relation between the explanatory factors (i.e. the parts) and the system to be explained (i.e. the whole) is represented as a spatial hierarchy. That is, they take into account only cases that satisfy not only Sarkar's criterion "fundamentalism" but also "spatial hierarchy". Hence, the first reading of 'fundamentality' I presented above (which equates Hüttemann's and Love's "fundamentality" with Sarkar's "fundamentalism") is not an option at all. But this is obscured by how Hüttemann and Love characterize the notion of fundamentality. Second, one important result of their analysis is that intrinsicality and fundamentality are two different aspects of reductive explanations, which can come apart. That is, a reductive explanation may fail to satisfy intrinsicality but, nevertheless, be reductive since it meets fundamentality (2011, 528-530). What is important to note is that this holds only if 'fundamentality' is understood as the qualified sense of fundamentality or if one assumes a certain notion of level, as I will explain in detail in Section 2.3.4.

Temporality

Hüttemann and Love emphasize a third aspect of reductive explanations in biology, namely temporality. They claim that the *temporal character* of reductive explanations (as well as their "intrinsic" character) has not received sufficient attention so far (2011, 521). According to their view, it is the temporal (and causal) character that distinguishes biological part-whole explanations from those in physics. Hüttemann and Love specify the notion of a temporal (causal) part-whole reduction as follows:

> Part-whole reductions (and explanations more generally) in biological science are often *temporal*. Properties of a whole at *t** are explained in terms of properties of parts at an earlier time *t*; the behavior of the parts at *t causes* the component to have a certain behavior or property at a later time *t**. (2011, 532)

This quotation reveals that, according to Hüttemann and Love, the temporal character of a reductive part-whole explanation consists in the temporal relation that exists between the properties mentioned in the explanans (i.e. the properties of the parts at *t*) and those cited in the explanandum (i.e. the properties of the whole at a later time *t**). For instance, the reductive explanation of muscle contraction exhibits a temporal and causal character since the property or state of the muscle fibers of being contracted at *t** is explained by the properties or states of its molecular parts (myosin, actin, tropomyosin, etc.) and the interactions between them at an earlier time *t*. In other words, it is shown how the property or state 'being contracted' of the muscle fiber at *t** is brought about (or caused) by the properties and interactions between myosin, actin, etc. at an earlier time *t*.[37]

2.3.3 The Reductivity of Part-Whole Explanations in Biology

What do these three aspects of reductive explanations that Hüttemann and Love identify imply for my central question of what it is that determines the reductive character of a biological explanation? It seems to me that the first thing to note is that one of these three aspects is not on par with other two. Temporality is not a *determinant* of the reductive character of an explanation, whereas intrinsicality and fundamentality are. It is true that, as Hüttemann and Love argue, most reductive part-whole explanations in biology possess a temporal character (and that this is an important fact to note)[38]. But this holds also for biological explanations that are non-reductive; many of them exhibit a temporal character, too. Thus, temporality does not help to distinguish reductive from non-reductive explanations. The opposite is true for intrinsicality and fundamentality. The fact that an explanation refers solely to intrinsic or to more fundamental factors is characteristic only for *reductive* explanations. Accordingly, intrinsicality and fundamentality are the features that determine the reductive character of biological explanations, whereas temporality is *not* such a feature. I do not want to accuse Hüttemann and Love of overlooking this difference, but of not having sufficiently emphasized it. Passages such as "explanations are *reductive*

---

[37] Hüttemann and Love regard the temporal nature of reductive explanations in biology as one reason why traditional concepts like identity, multiple realization, and supervenience (which are atemporal concepts) are of limited significance to biology (2011, 526, 533).

[38] This is important to note because, on the one hand, the temporal character distinguishes biological from physical part-whole explanations and, on the other hand, once part-whole relations are "treated temporally, intrinsicality and fundamentality take on independent significance in reductive explanations" (2011, 530).

explanations provided they conform to certain additional constraints, including intrinsicality, fundamentality, and temporality" (2011, 524) suggest that all three aspects serve to distinguish reductive from non-reductive explanations. But this is not true since the reductivity of a biological explanation only depends on whether it appeals exclusively to intrinsic factors or also mentions extrinsic factors and whether it solely refers to fundamental factors or also involves non-fundamental factors. By contrast, whether a biological explanation exhibits a temporal character or not does not as such affect its reductive character (although it affects whether intrinsicality and fundamentality can be met independently; see 2011, 528-530). Basically, this is also the conclusion Hüttemann and Love draw. But they fail to sufficiently stress the different status of the three aspects.

After having clarified this, let us now turn to the different ways in which biological explanations can succeed or fail to be reductive. Hüttemann and Love argue that

> there are two basic ways for a temporal part-whole reduction to fail as a *reductive* explanation…: either the behavior on an *extrinsic* element from the fundamental realm in necessary (intrinsicality fails…) or the behavior of an extrinsic system with a *non-fundamental* property is necessary (intrinsicality and fundamentality fail…). (2011, 539)

Put another way, temporal part-whole explanations can fail as reductive explanations either if intrinsicality fails or if intrinsicality and fundamentality fail.[39] Consider the examples Hüttemann and Love offer. The first case, the failure of intrinsicality, is illustrated by the explanation of how differential blood flow influences the shape of the heart during embryogenesis (Hove, et al. 2003). Since the blood cells referred to in the explanans – at least to the time they have an effect on the shape of the heart – are extrinsic to the heart (i.e. the system to be explained) intrinsicality is violated. However, it is important to note that according to Hüttemann and Love this failure of intrinsicality does not give rise to the non-reductive character of the explanation.[40] Rather, since all explanatory factors belong to "the qualified fundamental level of cells" (2011, 529) fundamentality is met and the explanation is said to be reductive. This shows that Hüttemann and Love treat intrinsicality and fundamentality not as necessary, but as *sufficient conditions* for the reductivity of an explanation. This is why they stress that the "success or failure of a reductive explanation is not an all or nothing phenomenon" (2011, 524). The second case, the failure of both

---

[39] Hüttemann and Love exclude the possibility that only fundamentality is violated, whereas intrinsicality is still satisfied (and, thus, the explanation can be characterized as reductive). They argue that this is not possible because "if a feature is intrinsic to *S* then in order to be contained within S it must be instantiated on a more fundamental level than *S* itself" (2011, fn. 12). As I will object in Section 2.3.4, this is true only if one interprets fundamentality according to Sarkar's "spatial hierarchy" criterion and not if one presupposes Hüttemann's and Love's qualified sense of fundamentality (see Section 2.3.2).

[40] One needs to be careful at this point since Hüttemann and Love claim that there is not just *one* reductive character of explanations. Rather, different respects in which an explanation can be reductive or not need to be distinguished. Accordingly, they would have done better to have introduced two notions of reductivity, namely reductivity$_F$ (for reductivity with respect to fundamentality) and reductivity$_{IF}$ (for reductivity with respect to intrinsicality and fundamentality). The explanation discussed above is only reductive$_F$, but not reductive$_{IF}$.

intrinsicality and fundamentality, is illustrated by the explanation of how direct interactions of different organs with the heart during embryogenesis change the hearts morphology. In this explanation both aspects of reductive explanations are violated because the explanation of the development of the heart appeals to other organs, that is, to factors that are extrinsic to the heart as well as located on the same spatial level than the heart (i.e. on a non-fundamental level). Since the explanation meets neither the intrinsicality criterion nor the fundamentality criterion it is characterized as non-reductive. Another example Hüttemann and Love discuss is the explanation of protein folding by reference to molecular chaperones, which provide the required environment for folding or which actively facilitate folding. According to their analysis, this explanation is non-reductive, too, because chaperones are extrinsic to the polypeptide as well as located on a higher (i.e. non-fundamental) spatial level than the polypeptide (2011, 538-541). However, as Hüttemann and Love frequently emphasize, such a failure of *reductivity* does not imply a failure of *explanation*. The question of whether a putative explanation is adequate (or not) must be distinguished from the question of whether an adequate explanation is reductive (or not). Recall that this distinction figured prominently in Sarkar's analysis, too (see Section 2.2.2).

Against this background the question arises why temporality is a relevant aspect of reductive explanations at all. The answer Hüttemann and Love give is the following: only if biological part-whole explanations are conceived as what they (in most cases) are, namely as *temporal* (respectively causal) explanations, does one become aware of the fact that intrinsicality and fundamentality are two *distinct* aspects of reductive explanations, which can be met independently.[41] Otherwise, what counts as fundamental is identified with what is intrinsic to the system of interest. I concur with Hüttemann and Love that intrinsicality and fundamentality are two distinct constraints of the reductive character of a biological explanation and that this is a significant difference, which has been overlooked by most philosophers of biology (but not by all; see e.g. Wimsatt 2006a, 2007). Accordingly, in developing my own account I will pick up this difference and elaborate it (see Chapter V). However, despite this general agreement I doubt that the temporal character of biological part-whole explanations is the only and the most important reason why fundamentality can be decoupled from intrinsicality, as I will explain in the subsequent section.

2.3.4 Shortcoming of Hüttemann's and Love's Account

All in all, the analysis of Hüttemann and Love provides several notable insights into the character of reductive explanations in biology as well as into the constraints that determine the success and failure of the reductivity of explanations. Some of these insights will be taken up and further elaborated in Chapter V. But their account also faces serious difficulties, which I reveal in this section. The first set of objections concerns Hüttemann's

---

[41] Strictly speaking, they claim that only fundamentality is decoupled from intrinsicality, not vice versa (see footnote 32).

and Love's notion of fundamentality, the second set of objections bears on their theses about temporality, and the last objection concerns the relation between reductive and part-whole relations.

First, as I have already demonstrated in Section 2.3.2, it remains unclear how exactly Hüttemann and Love specify the notion of fundamentality. I have identified three different readings of this notion: 'being more fundamental' can either mean what Sarkar refers to with his criterion "fundamentalism" (i.e. 'belonging to a more fundamental realm'), or it can be interpreted in line with Sarkar's criterion "spatial hierarchy" (i.e. as 'being located on a lower spatial level'), or it appeals to what Hüttemann and Love call a qualified sense of fundamentality (i.e. it means 'belonging to a restricted set of factors'). I have also called attention to the fact that the first reading contradicts Hüttemann's and Love's assumption that reductive explanations can be identified with part-whole explanations or, as they call them, part-whole reductions. The reason for this contradiction is that in biological part-whole explanations the system in question (i.e. the whole) is represented as being *spatially decomposed* into its parts. And if reductive explanations are identified with part-whole explanations, there are no reductive explanations that fulfill *only* Sarkar's "fundamentalism" (and not also "spatial hierarchy"). Thus, even if Hüttemann and Love seem to endorse "fundamentalism" as a possible reading of the notion of fundamentality, they can only consistently adopt the other two readings, namely spatial-hierarchy fundamentality and qualified fundamentality. But even if this has been clarified, at many points of their argumentation it still remains obscure which of these two notions Hüttemann and Love presuppose (although at some points they indicate when they refer to qualified fundamentality). Let me illustrate this by two examples.

The first example is their claim that fundamentality is decoupled from intrinsicality (i.e. that it is possible for temporal part-whole explanations to fulfill fundamentality but violate intrinsicality). Hüttemann and Love argue that this decoupling is made possible only by the temporal character of an explanation. Consider the explanation of how adrenergic hormones secreted in the environment of the heart can modulate heart rhythm. Hüttemann and Love claim that without this explanation being temporal it would be impossible that extrinsic (but fundamental) factors like hormones would have a causal influence on the heart and, thus, would be referred to in the explanation. I concur with Hüttemann and Love in this point. However, it seems to me that temporality is not what matters most when one tries to understand why an explanation can meet fundamentality but simultaneously violate intrinsicality. Rather, we need to understand what it means for an explanatory factor to fail to be a spatial part of the system in question (i.e. to be extrinsic) and, nevertheless to be more fundamental than the system. It is exactly at this point where the two different readings of the notion of fundamentality come into play. What I think is important to note is that a factor, which is extrinsic to a system, can only be more fundamental than the system according to the qualified sense of fundamentality, and not according to the spatial-hierarchy sense of fundamentality.[42] This is because factors that are located on a lower spatial level than the system are by definition spatial parts of the system (i.e. intrinsic to the system), whereas intrinsic as well as extrinsic factors can belong to the restricted set of factors, like in the hormone example the set of all macromolecules.

The second example, that reveals the importance of clarifying which notion of fundamentality is meant, is Hüttemann's and Love's claim that intrinsicality cannot be independently satisfied from fundamentality. They argue that this is the case because a factor that is located inside the system is automatically (qua being a part of the system) located on a lower spatial level than the level of the system (see 2011, fn. 12 and 22). However, Hüttemann and Love fail to point out that this is true only if fundamentality is understood as spatial-hierarchy fundamentality, and not as qualified fundamentality. For instance, imagine a qualified sense of fundamentality according to which only the restricted class of molecules count as fundamental. Based on this, the explanation of, let us say, protein synthesis, which refers also to non-molecular parts of the cell (like ribosomes or the nucleus membrane), satisfies intrinsicality (since it refers only to spatial parts within the cell), but violates this qualified sense of fundamentality. This would be a clear example, in which the success of intrinsicality is combined with the failure of (qualified) fundamentality.

Second, with regard to the aspect of temporality, different objections arise, too. To begin with, Hüttemann and Love do not sufficiently underline the different status temporality has with respect to intrinsicality and fundamentality. Contrary to these two aspects, temporality is not a determinant of the reductivity of explanations (see Section 2.3.3). Furthermore, it is true that previous analyses of epistemic reduction in biology have focused on atemporal concepts like multiple realization, identity, and supervenience. But this does not imply that the temporal and causal character of biological part-whole explanations has been overlooked until now, as Hüttemann and Love sometimes seem to suggest. For instance, Kauffman (1970) explicitly emphasizes the causal character of part-whole explanations (see Section 2.1.1). Likewise, the New Mechanists stress the importance of temporality for (mechanistic) part-whole explanations. They claim that the behavior of a whole system is explained by describing the entire *causal process* (i.e. mechanism) of how the interactions between the parts bring about the system's behavior. Moreover, they conceive the *temporal* organization of the parts of a mechanism, i.e. the rates, order, and duration of the stages of a mechanism, as crucial for an explanation (e.g. Craver/Darden 2001, 114f, 127f).

My main objection to Hüttemann's and Love's notion of temporality concerns the way they specify the temporal character of biological explanations. Recall that in their view the temporality of biological part-whole explanations traces back to the fact that they explain "a property of a whole at *t**… in terms of properties of its parts at an earlier time *t*" (2011, 531). What I think is misleading with this characterization is that it focuses exclusively

---

[42] At least if the notion of a spatial level is defined locally (see Chapter V).

on the temporal relation between the phenomenon to be explained (i.e. the properties of a whole at *t**) and the explanatory factors (i.e. the properties of the parts at an earlier time *t*). What Hüttemann and Love overlook is the fact that in most cases what is described in the explanans is itself a temporally extended process. In other words, in most biological (causal) explanations not the properties or states of parts at *one* time *t*, but the *entire process* of the interactions between the parts and the caused changes of their properties are cited as explanatorily relevant.[43] Consider the example of photosynthesis. In order to give a part-whole explanation of this phenomenon a highly extended temporal process starting with the absorption of light energy and finishing with the last step of the Calvin cycle is described. In many explanations what happens in the interim time is filled out, that is, how each stage of the process gives rise to the following stage, for instance, how the electrons flow down the electron transport chain and how this leads to the ultimate reduction of NADP to NADPH. Granted, some explanations do not entail the entire causal mechanism of photosynthesis, but rather assign only a choice of stages as explanatorily relevant. But the "explanatory factors" described in the explanans always involve temporal relations by themselves and this is crucial for the adequacy of these explanations. In their characterization of the temporal character of part-whole explanation in biological science Hüttemann and Love completely disregard this important kind of temporality.[44]

Finally, it remains unclear which relation between reductive explanations and part-whole explanations (or "part-whole reductions"; 2011, 527) Hüttemann and Love assume. The first impression one gets is that they equate these two kinds of explanation since they align themselves with authors that conceptualize reductionism in terms of the relationship between parts and wholes (2011, 527). But phrases such as "part-whole reductions can fail as reductive explanations" (2011, 528) suggest that there exist *part-whole reductions* that are non-reductive explanations (which is odd) or, at least, that there exist *part-whole explanations* that are non-reductive, namely those part-whole explanations that violate intrinsicality as well as fundamentality.[45] According to this interpretation, only a subset of part-whole explanations would be *reductive* part-whole explanations. However, this conflicts with Hüttemann's and Love's initial thesis that reductive explanations can be identified with part-whole explanations.

---

[43] This does not mean that in any part-whole explanation *each step* of the causal process from *t* to *t** is described. Sometimes there exist "gaps'" in the causal process that is cited as explanatorily relevant. These gaps are due to the unavailability of knowledge or to certain explanatory interests. Nevertheless, even in those cases the explanans is not restricted to the properties of the parts to one particular time *t*.

[44] I suppose that this is why Love and Hüttemann add the bracket "(or t_1, …, $t_n < t^*$)" to claims such as "the state or properties of the parts and their interactions at *t* (or $t_1, …, t_n < t^*$) bring about a change in the state or properties of the compound at time *t**" (2011, 188) in their second paper on "Comparing Part-Whole Reductive Explanations in Biology and Physics".

[45] The alternative would be to claim that explanations that violate intrinsicality and fundamentality, are neither reductive explanations nor part-whole explanations.

A more plausible view is that any part-whole explanation simultaneously is a reductive explanation. That is, whenever intrinsicality and fundamentality are both violated we have an explanation that exhibits *neither* a reductive *nor* a part-whole character. Furthermore, in cases in which intrinsicality is violated, but fundamentality is satisfied (as in the example of differential blood flow causing heart asymmetry) the explanation can be characterized as reductive (with respect to fundamentality). However, it is odd to call an explanation that refers to more than just to the parts of a system a "part-whole explanation". Thus, according to this view part-whole explanations are all reductive explanations, but there may be reductive explanations that are not part-whole explanations. These relations are illustrated by the following table.[46]

intrinsicality	fundamentality	reductive character	part-whole character
%	%	%	%
%	+	+	%
+	+	+	+

Table III.2: The Relation between Intrinsicality, Fundamentality, the Reductive, and the Part-whole Character of Biological Explanations

3 Interim Conclusion

In this chapter I critically discussed the different views on explanatory reduction in biology that have been proposed (or implicitly assumed) in the previous debate. The goal of this critical discussion was to identify adequate and fruitful insights concerning explanatory reduction, which can be included in my own analysis of explanatory reduction (see Chapter V), and to sort out those ideas about explanatory reduction that have proven to be unconvincing.

One of my central theses in this chapter is that there exist two different perspectives on explanatory reduction in the debate. The first perspective is the understanding of explanatory reduction that implicitly underlies Rosenberg's defense of explanatory reductionism (Section 1). He treats explanatory reduction as a relation between a higher-level and a lower-level explanation of the same phenomenon. Contrary to this approach, Sarkar and more recently Hüttemann and Love focus on individual explanations and examine the conditions under which they exhibit a reductive character (Section 2). I claimed that in the case of individual reductive explanations the relation of reduction exists between the two parts of an explanation, namely between the representation of the explanandum phenomenon and the representation of the explanatory relevant factors (explanans). Important precursors of this second perspective are Kauffman and Wimsatt.

---

[46] Hüttemann (personal communication) agrees to this view.

In Section 1.1 I introduced the major elements of Rosenberg's "Darwinian Reductionism". More specifically, I pointed out that Rosenberg has not always been a defender of reductionism (although he has always been inclined to it), and that his recent return to reductionism is due to his abandonment of Nagel's model of theory reduction (Section 1.1.1). I argued that the version of explanatory reductionism Rosenberg advocates is quite radical: he presents his Darwinian reductionism as being non-eliminative (although it remains unclear what exactly this amounts to), as having a broad scope (i.e. as applying to all biological fields and all biological phenomena), and as having a restricted locus (i.e. he claims that there is just a single reducing level, namely the level of molecules) (Section 1.1.2). Finally I presented the core arguments Rosenberg offers in favor of his Darwinian reductionism (Section 1.1.3). Briefly speaking, his argumentation runs as follows: Functional biology yields ultimate explanations as well as proximate explanations, which are also implicitly ultimate. Functional biology's ultimate explanations are only incomplete how-possible explanations. Because in the context of advanced biological inquiry only maximally complete explanations are adequate, functional biology's ultimate how-possible explanations need to be turned into (i.e. reduced to) molecular proximate why-necessary explanations.

On the basis of this information I could then consider the question what Rosenberg's view of explanatory reduction is (Section 1.2). The units of reduction were quickly identified. Rosenberg treats explanatory reduction as a relation between an ultimate how-possible and a molecular proximate why-necessary explanation of the same phenomenon. I further argued that the process of reduction he envisages is characterized as the transformation of the ultimate explanation into the proximate explanation in molecular terms, which involves elements of completion and correction. However, I stressed that it remains obscure whether some parts of the ultimate explanation are maintained during this transformation process and what, exactly, the result of this process looks like.

Finally, I revealed some serious shortcomings of Rosenberg's perspective on explanatory reduction (Section 1.3). I argued that he fails to detach his account from the Nagelian framework and that his perspective on explanatory reduction is shaped too much by the philosophical requirements his defense of reductionism carries with it. This is why Rosenberg's account fails to capture what epistemic reduction and what explanation in contemporary biological practice really is. Furthermore, I claimed that Rosenberg draws an artificial line between questions of reduction and questions of explanation that cannot be sustained – not even on the basis of his own argumentation (see also Chapter IV, Section 2.2). These drawbacks led me to dismiss Rosenberg's perspective on explanatory reduction.

In Section 2 I critically examined the work of Kauffmann, Wimsatt, Sarkar, Hüttemann, and Love. What binds these authors together is that they focus on individual reductive explanations and analyze the constraints of their reductivity. Thus, I take them to provide an alternative, second perspective on explanatory reduction according to which reduction is not a relation between two explanations, but a feature of an individual explanation (respectively the relation between the explanandum and explanans of an explanation). Most notably, this perspective has been put forward by Sarkar and, more recently, by Hüttemann and Love. In what follows I summarize what I regard as their most important findings and the respects in which I think their analysis is deficient or implausible. This summary is instrumentalist because it condenses only those issues that can be utilized for developing my own account of explanatory reduction in Chapter V.

First, one result of my examination is that reductive explanations are closely related to other kinds of explanation. In fact, Wimsatt identifies reductive explanation with mechanistic explanation (Section 2.1.2), Hüttemann and Love equate reductive explanation with part-whole explanation (Section 2.3.3), and Sarkar, Hüttemann, and Love argue that reductive explanations are to be characterized as lower-level explanations (Sections 2.2.3 and 2.3.2). I agree that these four kinds of explanation are closely connected. But I also think that it is important not to identify reductive explanation with mechanistic, part-whole, or even lower-level explanation. My analysis will show that although the set of mechanistic, part-whole, and lower-level explanations overlap the set of reductive explanations to a large extent, they do not coincide. And I will argue that it is crucial to know in which respects they do not coincide.

Second, Sarkar argues that an appropriate analysis of reductive explanation should focus on substantive issues, not on formal ones (Section 2.2.1). This is a crucial insight that stands behind much of the criticism that can be raised against Nagelian models of theory reduction (see Chapter II, Section 4). In accordance with Sarkar's focus on substantive issues, I develop an account of explanatory reduction that can be characterized as "ontic" because it traces the reductive character of an explanation back to the fact that it appeals to certain relations and facts that exist in the world (see Chapter V, Section 6).

Third, the major part of Sarkar's, Hüttemann's, and Love's analysis is that they highlight different aspects of reductive explanations in the biological sciences, respectively propose different criteria of reductivity. That is, they provide an answer to the central question I am concerned with, too, namely what are the features that reductive explanations have in common (and according to which different types of reductive explanations can be distinguished). Sarkar identifies three criteria of reductivity, namely fundamentalism, abstract hierarchy, and spatial hierarchy (Section 2.2.3). Hüttemann and Love distinguish three aspects of reductive explanations, namely intrinsicality, fundamentality, and temporality (Sections 2.3.2 and 2.3.3).

At this point, let me just briefly sum up my three central points of criticism (Sections 2.2.4 and 2.3.4). First, with respect to Sarkar's account I objected that his focus on molecular biology gives rise to a too-narrow interpretation of the "spatial hierarchy" criterion. In my words, he identifies lower-level explanation with molecular explanation (i.e. fundamental-level explanation). In Chapter V, Section 2.3.1 I will show that fundamental-level explanations are just a subtype of lower-level explanations. Hüttemann and Love seem to adopt a broader notion of what they call "fundamentality". However, I claimed that it remains too obscure how exactly they specify the notion of fundamentality. Second, Sarkar presents his criterion "fundamentalism" as a discrete, sufficient criterion that distinguishes a significant class of reductive explanations. I argued that this is an unconvincing assumption because the mere fact that some factors are included in an explanation and others are ignored is not sufficient to render an explanation reductive. Third, Hüttemann and Love highlight "temporality" as a third aspect of reductive explanation. However, I claimed that they fail to emphasize the different status temporality has with respect to intrinsicality and fundamentality. That is, contrary to the latter two, temporality is not a determinant of the reductivity of explanations. Moreover, I showed that Hüttemann and Love misrepresent the temporal character of biological explanations since they fail to represent the explanans itself as a temporal process.

Besides these problematic theses Sarkar's, Hüttemann's, and Love's accounts include valuable insights as well. For instance, the criteria or aspects spatial hierarchy and fundamentality express is the important finding that in reductive explanations the factors that are referred to in the explanans are located on a lower level than the phenomenon to be explained (Sections 2.2.3 and 2.3.2). Furthermore, Hüttemann and Love are the first authors to recognize a difference between the claim that something is located on a lower level (fundamentalism) and that something is located inside the spatial boundary of a system (intrinsicality) (Section 2.3.2). This difference will be taken up and further elaborated in my own analysis of reductive explanation in biology.

To conclude, the second perspective, which focuses on individual reductive explanations and analyzes the conditions under which explanations succeed and fail to be reductive, seems to be the much more promising way to analyze explanatory reduction than Rosenberg's perspective, which treats reduction as a relation between an ultimate how-possible and a molecular proximate why-necessary explanation of the same phenomenon.[47]

---

[47] It might be that the inappropriateness of the second perspective primarily traces back to the peculiarities of Rosenberg's view. That is, I do not preclude that it might be in principle a valuable undertaking to reconstruct explanatory reduction as a relation between different explanations, rather than as a relation between explanandum and explanans of an individual explanation. It appears to me that such an approach is particularly promising when we are concerned with successional explanatory reduction, and not with interlevel explanatory reduction.

IV A Closer Look at Biological Explanations

> *"[S]cientists nowadays increasingly question the validity of reductive explanations… The debate between reductionists and antireductionists is thus very much a debate about what constitutes a good scientific explanation."*
>
> (Marc H. V. van Regenmortel 2004a, 145)

> *"[T]he debate [about reductionism] cannot be a dispute about 'explanation', for example a disagreement about pragmatic, erotetic, Protagorean versus nonerotetic accounts of explanation. For that is a general problem in the philosophy of science, not a problem about reductionism in the philosophy of biology."*
>
> (Alex Rosenberg 2006, 41, fn. 6)

Content

If one decides to examine epistemic reduction by focusing on reductive explanations (as I do) the question arises how entangled the issue of reduction becomes with the issue of explanation. With respect to reductionism the two quotations above express contradictory stances on this question. van Regenmortel, a molecular biologist who is very interested in philosophical discussions about reductionism (e.g. van Regenmortel/Hull 2002), claims that the debate about reductionism is, basically, a debate about what constitutes a good (or an adequate) explanation. Rosenberg opposes this view. He stresses that the debate about reductionism does not amount to a discussion about which of the competing views about explanation (each of which identifies different features that adequate or good explanations must possess) is correct. As I argue in Section 2.2, discussions about whether explanatory reductionism is true or not in fact amount to specific questions about scientific explanation.

However, recall that the target of my analysis is not reductionism, but reduction (see Chapter II, Section 1). My main goal in this chapter is thus to reveal the interrelations that exist between the issue of explanation and my account of explanatory reduction. For instance, how neutral is my account with respect to different models of explanation? Does the "ontic" character of my account of reduction commit me to an ontic account of explanation? Which stance on the pragmatics of explanation is most consistent with my account? What distinguishes higher-level from lower-level explanations and is there an "objective" reason to prefer lower-level explanations? Questions about explanation that are involved in disputes about explanatory reductionism will be of interest, too – but only insofar as they concern issues that impact my account of reduction. All in all, the choice of topics addressed in this chapter is quite instrumentalist. I approach only those issues that are relevant to my project of developing an ontic account of explanatory reduction. Not only is the set of questions I approach selective, but the set of answers I give is selective, too. As we will see, at some points it is necessary to take a stand and to argue for a certain view with respect to explanation. But with regard to other decisions I claim that it is better to abstain since it renders my account of explanatory reduction neutral with respect to different positions in the debate about explanation.

Many philosophical discussions about the concept of explanation have focused on the central question of what the nature of (scientific) explanation is. Which feature is common to all explanations and distinguishes them from non-explanatory kinds of scientific achievements (i.e. primarily from pure descriptions; possibly also from purely predictive models, mere sketches, models including explanatorily irrelevant factors; see Craver forthcoming a)? In other words, where does the explanatory force in explanations stem from? At the outset of this chapter I give a brief overview on the different answers that have been proposed to this question (Section 1). In particular, I discuss two major kinds of accounts of explanation: first, the *covering-law (CL) model* of explanation (Section 1.1), under which I subsume Hempel's deductive-nomological (DN) model, his inductive-statistical (IS) model (Hempel/ Oppenheim 1948; Hempel 1965), as well as the unificationist account of explanation advocated by Michael Friedman (1974) and Kitcher (1981, 1989, 1999b);[1] second, the *causal-mechanical (CM) model* of explanation (Section 1.2), which has been introduced and refined by Wesley Salmon (1984a, 1994, 1997) and further developed by the New Mechanists (e.g. Machamer, et al. 2000; Glennan 2002; Bechtel/Abrahamsen 2005; Craver 2007a) and by proponents of an interventionist theory of causation (e.g. Woodward 2003; Strevens 2008). This section serves to introduce these different accounts of explanation as well as the criticism that has been offered against them. Since this is both an extensive and a well-known debate I will restrict my overview to the main characteristics of these accounts. In Section 1.3 I address in more detail a question that is of particular interest to my discussion about reduction, namely what it means when proponents of the CM model call their conception of explanation "ontic" (e.g. Salmon 1984b, 296; Craver forthcoming a). This is important to me since I characterize my own account of explanatory reduction as ontic, too, and, in doing so, draw on how this term is understood in the debate about explanation. In this section I point out that there are two different meanings of 'ontic' circulating in the debate and that my usage of this concept is confined to just one of them.

The introductory notes on CL and CM models of explanation provide the basis on which I can address more specific questions about explanation that concern the issue of reduction in Section 2. The point from where I start is the already introduced dispute between those who argue that the debate about explanatory reduction amounts to a debate about explanation (e.g. van Regenmortel) and those who deny this (e.g. Rosenberg). The truth of these assertions must be judged separately with regard to debates about the adequate account of reduction (Section 2.1) and with regard to debates about the correctness of reductionism (Section 2.2).

In Section 2.1.1 I point out why the question of what constitutes the reductive character of biological explanations (the question of reduction) does not boil down to the question of what characterizes an adequate explanation (the question of explanation). In line with this, I argue that my account of explanatory reduction remains uncommitted with respect to whether the CL or the CM model adequately captures the nature of explanation, and that this is a good thing, too. Section 2.1.2 clarifies some terminological issues concerning my way of speaking about biological explanations.

In Section 2.2 I turn to the question whether and in which respect debates about explanatory reductionism depend on discussions about explanation. I begin by examining what these discussions about explanation are to which questions about reductionism allegedly amount to (Section 2.2.1). The result is that those discussions about explanation that matter to reductionism concern questions about the adequacy of higher- and lower-level explanations. In Section 2.2.2 I argue that disputes about reductionism in fact amount to answering these specific questions of explanation. The answers one gives to these questions seem to be highly affected by one's stance on the pragmatics of explanation

---

[1] I give reasons for this classification below (see Section 1.1).

(rather than by one's stance on whether CL or CM models are adequate). Hence, in Section 2.2.3 I address the question whether and in which way pragmatic factors influence the adequacy of explanations. After reviewing van Fraassen's pragmatic account (1977, 1980) I consider Rosenberg's criticism that pragmatic accounts of explanation are misleading because they turn the adequacy of an explanation into a "subjective" matter of what the background knowledge and explanatory interests of biologists are. I reject Rosenberg's criticism by pointing out how van Fraassen's account can be refined. This refinement enables me, on the one hand, to claim that the difference between an adequate explanation and a non-explanation is "perfectly objective" (Sober 1999, 550) and, on the other hand, to admit that whether or not a factor is explanatorily relevant depends on pragmatic factors (since the choice of the relevance relation is context-dependent). Against this background I point out why there cannot be just one right level of explanation, irrespectively of the research context in which an explanation is developed (Section 2.2.4).

The issues I approach in Section 2.2 occupy center stage in the long-standing debate about the truth of explanatory reductionism. However, despite their centrality to the reductionism dispute one might question whether these issues are relevant to my work of seeking an adequate account of explanatory reduction. To put it briefly, these issues are relevant, although their relevance is only indirect. Rosenberg's objection to pragmatic accounts of explanation must be overruled since his special view of reductive explanation restricts the set of "empirical data", on which my analysis is built, in an inappropriate way. If Rosenberg were right only molecular explanations would count as reductive explanations. But the explanatory practice in biology shows a much larger variety of reductive explanations (see Chapter V), which needs to be taken into account. The revised version of van Fraassen's pragmatic account of explanation that I develop in Section 2.2.3 ensures that all different kinds of explanation that can be found in actual biological practice are considered. Hence, I need the discussions presented in Section 2.2 in order to reject views of explanation that are incompatible with my account of explanatory reduction.

1 Accounts of Explanation

It is common to both main kinds of accounts of explanation that they seek to elucidate the nature of explanation. That is, both accounts aim at identifying those features that all explanations share (i.e. the "essential characteristics" Hempel/Oppenheim 1948, 135) and that distinguishes them from mere descriptions (or, more generally, from non-explanations or inadequate explanations[2]). In other words, they offer an answer to the question what it is

---

[2] I use the term 'explanation' as a success term. Thus, properly speaking, inadequate, failed or unsuccessful explanations do not exist. If an explanation of a phenomenon turns out to be inadequate or fails to succeed the putative explanation proves *not* to be an explanation *at all*. Note that this is compatible with the thesis that explanations can vary in quality. Explanations can be worse, but still be explanations. However, since avoiding phrases such as 'adequate explanation' would complicate the way of speaking, I sometimes stick to them.

to achieve an understanding of *why* something happens (as opposed to merely knowing *that* something happens).[3] Merely the answer they give differs. Proponents of a CL model argue that phenomena are explained by deriving them from certain laws and initial conditions and that the explanatory force arises from the fact that explanations render the phenomenon to be explained *nomically expectable* (respectively, that they subsume the phenomenon under a general pattern). Those who advocate a CM model disagree. They claim that explanations make phenomena understandable since they trace the causes that lead to or make up the phenomenon. That is, according to the CM model the explanatory force stems from the fact that explanations show how a phenomenon is situated in the *causal structure* of the world. Let us consider some details of these accounts and the objections they encounter.

1.1 Covering-Law (CL) Model

1.1.1 Varieties of CL Models

Hempel and Oppenheim identify two major constituents of an explanation, namely the *explanandum*, which is the sentence describing the phenomenon to be explained, and the *explanans*, that is, "the class of those sentences which are adduced to account for the phenomenon" (1948, 137). This terminology has gained broad acceptance in the debate about explanation, although some proponents of the ontic account challenge the assumption that the explanans is an epistemic entity (i.e. a statement or representation) and not something in the world itself (e.g. a fact or certain causes) (Salmon 1992, 10; Craver 2007a, 27, forthcoming a; see Section 1.3).

According to Hempel's *deductive-nomological (DN) model* (Hempel/Oppenheim 1948; Hempel 1965) explanations are sound deductive arguments, in which the explanandum is logically deduced from a set of premises (the explanans) containing at least one general law[4] and certain statements of antecedent conditions. Furthermore, the sentences constituting the explanans must have empirical content and be true. This general structure of explanation can be illustrated as follows:

$$\left.\begin{array}{ll} C_1, C_2, \ldots, C_n & \text{(statements of antecedent conditions)} \\ \underline{L_1, L_2, \ldots, L_n} & \underline{\text{(general laws)}\qquad\qquad\qquad\qquad} \end{array}\right\} \text{explanans}$$

$$\left.\begin{array}{ll} E & \text{(description of the phenomenon} \\ & \text{to be explained)} \end{array}\right\} \text{explanandum}$$

---

[3] Although I agree that, broadly conceived, explanations are answers to explanation-seeking *why-questions*, one should notice that explanations can be answers to other kinds of questions, too – for instance to how-questions, what-for-questions, etc. (e.g. Beatty 1990, 203; Dupré 1993, 106). These different kinds of questions point to different types of explanation (see Section 2.2.2).

[4] Hempel uses the term 'law' as an epistemic notion. That is, he refers with it to law statements. In this section I will adopt this way of speaking. However, in the contemporary debate it has become established to speak of law statements and to use the term law as an ontological notion. In the rest of my book I therefore assume this more common reading.

The DN model applies to the explanation of particular phenomena as well as to the explanation of generalizations (respectively laws). The basic idea behind this model is that a phenomenon is explained by showing that, given the particular circumstances and the laws in question, the phenomenon was to be expected. For instance, the temporary drop of the mercury column in a thermometer and its subsequent swift rise is shown to be expected in light of certain antecedent conditions (e.g. that the thermometer consists of a glass tube which is partly filled with mercury, that it was immersed into hot water) and certain general laws (e.g. the laws of thermic expansion of mercury and of glass). Thus, according to the DN model the nature of scientific explanation can be described as "*nomic expectability*" (Salmon 1989, 57). Because the general laws cited in the explanans typically "cover" the phenomenon to be explained (since it is an instance of the law), the DN model is also referred to as the *covering law (CL) account* of explanation.

If the explanans includes statistical rather than deterministic laws we have a case of a statistical explanation. Hempel (1965) distinguishes two types of statistical explanations. According to the deductive-statistical (DS) model statistical regularities are explained by deduction from more comprehensive statistical laws, whereas according to the inductive-statistical (IS) model particular phenomena are explained by subsuming them under statistical laws. IS explanations deviate from DN and DS explanations in an important respect: they are "inductive" rather than deductive. For an IS explanation to be adequate Hempel requires that the explanans confers high probability on the explanandum.[5]

The *unificationist account* of explanation that was introduced by Friedman (1974) and elaborated by Kitcher (1981, 1989, 1999b) is often discussed as a third major account of explanation. However, Kitcher himself emphasizes the affinity between his account and Hempel's CL model. In particular, he characterizes the unificationist account as an explication of the "unofficial view" (Kitcher 1981, 508) of explanation that stands behind Hempel's official model. What is more, according to Kitcher, explanation is a matter of *deductively deriving* descriptions of many different phenomena by using as few and as stringent[6] "argument patterns"[7] (1981, 515) as possible. In other words, Kitcher regards explanations as a certain kind of DN explanations, namely as such arguments that instantiate an argument pattern that belongs to the "explanatory store" (1981, 512), that is to the set of argument patterns that maximally unifies the set of beliefs accepted at a particular time in science. I therefore classify the unificationist account as an advanced version of Hempel's DN

---

[5] The issue of statistical explanation will not receive extensive treatment here. For more details about Hempel's analysis of statistical explanation and about alternative approaches (as e.g. Salmon's Statistical relevance (SR) model) see for instance Woodward 2011; Salmon 1984a, 1989.

[6] Those argument patterns are said to be stringent that contain some non-logical expressions and that are fairly similar in terms of logical structure (Kitcher 1981, 518, 527)

[7] An argument pattern is an ordered triplet consisting of a schematic argument, a set of sets of filling instructions (one for each term of the schematic argument), and a classification of the schematic argument (Kitcher 1981, 516).

model and, thus, as a subtype of the CL account.[8] However, this does not imply that there exist no differences between the DN model and Kitcher's account. For instance, they even differ with regard to what they regard as most important for the nature of explanations. The DN model highlights nomic expectability, whereas the unificationist account treats the relation between explanans and explanandum to be primarily a relation of "*pattern subsumption*" (Strevens 2008, 13; my emphasis): the explanatory force arises from showing that the explanandum is an instance of a more general pattern. But on closer inspection it turns out that this is merely a difference in emphasis. Hempel's DN model can be interpreted as identifying pattern subsumption (in the form of "general-law subsumption") as a crucial characteristic of explanation, too.[9]

1.1.2 Objections to DN and IS Models

Since Hempel and Oppenheim have published their paper on "Studies in the Logic of Explanation" in 1948 many philosophers have discussed the virtues and shortcomings of the CL model. Today, several objections and counterexamples are well known. Thus, in what follows I will not discuss these objections in detail, but rather give a brief overview of them. This proceeding is appropriate because there is no need for me to take a stand on which model of explanation is most convincing (more on this in Section 2.1). At least five different objections to the DN model can be distinguished. The former two show that it fails to state *sufficient* conditions for adequate explanation. The third is a special problem the IS model faces, the fourth objection reveals that the DN model does not identify *necessary* conditions for something to be an explanation, and the final objection is of a general nature.[10]

First, cases like the flagpole example reveal that many explanations exhibit *asymmetric* or *directional features*, to which the DN model is insensitive. One can derive the length (*s*) of the shadow cast by a flagpole from its height (*h*), the angle (Θ) it makes with the sun, and certain laws describing the rectilinear propagation of light. But one can equally derive the height of the flagpole (*h*) from the length (*s*) of the shadow, the angle (Θ), and the same laws. Thus, according to the DN model not only the height of the flagpole (plus laws and antecedent conditions) explains the length of the shadow, but also vice versa. But we typically accept the former, but not the latter, as an explanation (since the flagpole together with the sun causes the shadow, but the shadow does not cause the flagpole). Thus, several authors have argued that the DN model fails to state sufficient conditions for explanation.

---

[8] Woodward (2003, 152-161) agrees with me on this point.

[9] This interpretation is confirmed by certain statements that Hempel makes. For instance, he claims that the understanding that explanation conveys lies "in the insight that the explanandum fits into, or can be subsumed under, a system of uniformities represented by empirical laws or theoretical principles" (1965, 488).

[10] Kitcher's unificationist account avoids the former two objections since it specifies *additional* criteria an explanation needs to satisfy in order to count as an explanation (i.e. that explanations must instantiate an argument pattern that belongs to the explanatory store that maximally unifies all accepted beliefs). However, it encounters new problems (compare, e.g., Craver 2007a, Chapter 2 and Woodward 2011, Section 5) and is still vulnerable to the fourth and fifth criticism.

Second, another kind of example shows that the DN model fails to sort out descriptions, which include explanatorily *irrelevant* information, as non-explanations. For instance, from the law “All males who take birth control pills regularly fail to get pregnant” and the antecedent conditions that John Jones is a male and has been taking birth control pills regularly one can deduce the sentence “John Jones fails to get pregnant”.[11] But we would refrain from calling this an explanation since the fact that John Jones takes birth control pills regularly is causally irrelevant for his childlessness. This kind of counterexample also reveals that the DN model does not provide sufficient conditions for explanation.

Third, the IS model has been opposed on the grounds that it is not necessary for a statistical explanation that the explanans shows that the explanandum was expected with *high probability*. For example, we would explain the fact that a person catches malaria by reference to the antecedent conditions that this person has been stung by a malaria mosquito and to the (imagined) statistical law “One percent of all people who has been bitten by a malaria mosquito catch malaria” even though the explanans does not confer high probability on the explanandum.[12]

Fourth, the requirement that all explanations *must contain laws* seems to be too strong and, thus, not to be a necessary condition for explanation. It is argued that several explanations in everyday life as well as in science do not explicitly refer to laws.[13] For instance, the phenomenon that the inkwell on my desk tipped over and fell to the floor can be explained by reference to the impact my elbow had on the inkwell, even though this explanation lacks an explicit DN structure (Scriven 1959). There are two possible responses to examples of this kind. Either one can simply deny that descriptions like “my knee hit the desk and caused the inkwell to tip over” are explanatory (which is implausible given the widespread acceptance of descriptions of this kind as explanations). Or one can adopt the so called “hidden structure strategy” (Woodward 2003, 159; 2011, Section 2.6). In pursuance of this strategy Hempel (1965) argues that explanations of particular phenomena sometimes are merely partial or elliptical explanations (that do not involve law statements), but that in all these cases there exists an underlying explanation, which makes explicit reference to laws. According to this line of reasoning, examples of law-free explanations (like the above explanation of why the inkwell tipped over) are not counterexamples to the DN model since even they refer to laws – with the difference that their reference to laws is merely “implicit” and, thus, their DN structure is “hidden”. Railton pursues a similar strategy. He states that the explanations we actually give are mostly incomplete and non-ideal since they convey only partial “explanatory information” (1981, 240) about why a certain phenomenon

---

[11] This example goes back to Salmon 1971, 34.

[12] As a response to these kinds of objections Salmon (1971, 1990) has developed his statistical relevance (SR) model, according to which statistical explanations must cite only statistically relevant factors. The SR model places no restrictions on the size of the probabilities.

[13] This is an objection that appears to be even more forceful which respect to biology since it is questionable whether there exist biological laws (at least in the strict sense) at all.

occurred. He contrasts them with the "ideal explanatory text" (1981, 246), which contains all of the causal and nomological information that is relevant to the phenomenon to be explained.[14] But Railton does not go so far to claim that only ideal texts constitute adequate explanations. Instead, he recognizes a "continuum of explanatoriness" (1981, 240). All in all, if the hidden structure strategy were successful it would invalidate the criticism that several accepted explanations do not refer to laws. However, the hidden structure strategy is confronted with serious problems.[15]

Fifth, the DN model faces the criticism that it relies on a concept which has not been sufficiently specified by itself, namely on the concept of a law of nature. Even today there is no consensus about how to define what a law is, that is, what the features of lawhood are, and how law statements can be distinguished from mere accidental generalizations (e.g. from the general statement "All birds on my balcony are titmice").

In sum, many of the criticisms that have been directed against the DN model reveal that "causality is a major focus" (Salmon 1992, 34). The reason why we think that the length of the shadow fails to explain the height of the flagpole and that John Jones' taking birth control pills regularly does not explain his childlessness is that the former are not *causes* of the latter. Likewise, counterexamples that are subject to the fourth objection show that one can explain a phenomenon by describing its causes – regardless of whether this description involves law statements or not. This observation led several authors to adopt an alternative view of explanation, which regards causality as being intimately involved in explanation.

1.2 Causal-Mechanical (CM) Model

The CM model of explanation was originally developed and further refined by Salmon (1984a, 1994, 1997). In more recent years, alternative versions of the CM model have been proposed. For instance, the New Mechanists have taken up the idea that to explain a phenomenon is to describe the causal mechanism that produces it (e.g. Machamer, et al. 2000; Glennan 2002; Bechtel/Abrahamsen 2005; Craver 2007a, forthcoming a). By contrast, other philosophers have focused on the more general thesis that explanations trace the causes of phenomena. They have tried to specify the notion of a causal explanation, for example by appealing to an interventionist theory of causation (e.g. Woodward 2003, Strevens 2008). All these accounts differ in their details, that is, how they specify the notion

---

[14] Railton deviates from Hempel in an important respect: he pursues the idea that explanations are not purely arguments but "accounts", that is, accounts in which arguments (and laws) play a central role, but "do not tell the whole story" (1981, 236). Railton argues that the ideal explanatory text does not only include derivations from laws, but also "attempts to *elucidate the mechanisms* at work" (1981, 242).

[15] For instance, the notion of "underlying" is too vague so that there are not sufficient restrictions on what counts as an underlying explanation. Furthermore, underlying ideal explanations are usually unknown or even epistemically inaccessible. This is problematic since it cuts the close connection between explanation and understanding. The underlying ideal texts would count as explanations even though they do not provide understanding (see also Woodward 2011, Section 2.6).

of a cause, which understanding of a causal mechanism they presuppose, which role they assign to the description of mechanisms in causal explanations, how “ontic” their account of causal explanation is, and whether and how they spell out the concept of explanatory relevance. Despite these differences all accounts share a *basic idea*, namely that to explain a phenomenon means to situate this phenomenon within the causal structure of the world. According to the proponents of a CM model neither the logical structure of a putative explanation nor the question whether the explanans renders the explanandum nomically expectable matters. Solely the question of whether the causes that lead to or make up the phenomenon of concern are adequately represented is of interest to whether the representation is explanatory or not. In order to illustrate this basic idea I give a brief overview of Salmon’s CM model, of the mechanistic account of explanation, and of the general criticism that CM accounts of explanation encounter.

1.2.1 Salmon’s CM Model

Initially Salmon attempted to characterize explanation in purely statistical terms (see his statistical relevance (SR) model; 1971). However, in 1984 he abandoned this attempt and advocated a CM view of explanation, which he called “the ontic conception”[16] (1984b, 296) and which he contrasted with epistemic and modal conceptions of explanation. Salmon’s CM model rests on a certain view of causation, namely that causation involves objects coming into spatio-temporal contact and transmitting something between them. Salmon developed two different versions of this *transmission theory*, the mark-transmission account (1984a) and the conserved-quantity account (1994; see also Dowe 1992, 2000).[17] But the key concepts of his account remained the same: causal processes (distinguished from pseudo-processes) and causal interactions.

According to Salmon’s earlier CM model *causal processes* possess the ability to transmit a mark in a spatio-temporally continuous way. That is, if a causal process is altered (i.e. marked) in a certain way this alteration will persist (i.e. be transmitted) from one spatiotemporal location to another. For instance, a bike whose lamp is broken will retain this mark even if it is moved. Salmon contrasts causal processes with pseudo-processes that lack the ability to transmit a mark (e.g. the shadow of a moving bike). Salmon characterizes *causal interactions* as a spatio-temporal intersection between two causal processes that modifies the structure of both. For example, a collision between two bikes that dents the front wheel of both bikes is a case of a causal interaction. In 1994 Salmon changed his CM model. He no longer treated the transmission of marks as being central for causation, but the transmission of conserved quantities (see also 1997, 1998). Accordingly, he defined a causal process as a world-line of an object that exhibits a non-zero amount of a conserved quantity (e.g. mass, energy, charge). A causal interaction is then an intersection of at least

---

[16] The notion of an ontic conception of explanation will be analyzed in Section 1.3.

[17] The primary motivation for this change was Salmon’s concerns about counterfactuals (1994).

two of these world-lines that involves the exchange of a conserved quantity. This is only a very rough overview of Salmon's account but it suffices for my concerns.

An aspect of Salmon's CM model, which is of particular importance for my purposes, is his distinction between the "etiological" and the "constitutive aspect" (1984a, 9) of causal explanation. In his view a certain phenomenon can be explained, on the one hand, by tracing the causal processes and interactions that lead up to it (the *etiological aspect*) and, on the other hand, by describing the causal processes and interactions that make up the phenomenon itself (the *constitutive aspect*). Both of these aspects are explanatory since they show how a phenomenon is embedded in its causal network – either by tracing its external causal network or by displaying its internal causal structure. I will argue below that the mechanistic conception of explanation takes into account only constitutive causal explanations (see Section 1.2.2).[18] Salmon also uses the notion of a causal mechanism to characterize his ontic conception of explanation. For instance, he writes that "explanatory knowledge is knowledge of causal mechanisms… that produce the phenomena with which we are concerned" (1989, 128). However, it is crucial to note that Salmon employs a very broad and unrestricted concept of a causal mechanism, according to which all causal processes and causal interactions count as causal mechanisms (1984b, 297). Thus, mechanistic explanations à la Salmon are not subtypes of causal explanations (but identical to them) since they are not restricted to the class of constitutive causal explanations.

Salmon's CM model suffers from serious limitations.[19] For reasons of space, I focus on two of them that are most relevant to the issue of reduction. First, it can be objected that although Salmon's CM model suffices to distinguish between causal processes and non-causal processes, it fails to tell us which features of a causal process are explanatorily relevant to the phenomenon to be explained. This is true for Salmon's mark-transmission account as well as for his recent conserved-quantity account. Recall the example of John Jones taking birth control pills. As Hitchcock (1995) recognizes there are spatio-temporally continuous causal processes (that transmit marks or exchange conserved quantities) and causal interactions at work when John Jones ingests birth control pills (e.g. the pills dissolves, certain constituents are transported into the blood, etc.[20]). Likewise, when Ms. Jones takes the pills spatio-temporally continuous causal processes and causal interactions are at work, too. Granted, they are different from those taking place in John Jones. And this difference is important because the causal processes and interactions in Ms. Jones are causally and explanatorily relevant for her childlessness, whereas the causal processes and interactions in

---

[18] For instance, Craver emphasizes that his goal is to "construct a normatively adequate *mechanistic* model of constitutive explanation (henceforth, mechanistic explanation)" (2007a, 111).

[19] Other objections concern, for instance, causation by omission and prevention. For an overview see Kitcher 1989; Hitchcock 1995; Woodward 2011, Section 4; Craver 2007a, Chapter 3, Section 3.

[20] Please notice that these processes must be described differently in the framework of Salmon's conserved-quantity CM model, namely as transmitting and exchanging conserved quantities.

John Jones are *not* causally and explanatorily relevant for his childlessness. But Salmon's CM model alone fails to capture this difference in relevance.[21]

Second, an objection that applies particularly to Salmon's actual account (respectively to Dowe's account of causation that Salmon takes for granted) is that it regards causal relations on higher levels to be dependent on the causal connections between their fundamental physical constituents. That is, in order to trace the causal processes and interactions that lead up to or make up a particular phenomenon to be explained we need to go down to the lowest, physical level of conserved quantities (since only on this level can we decide whether something is a causal process or not). One can object that with regard to several phenomena (especially those studied in higher-level sciences) this strategy seems to "lead us away from the right level of description" (Woodward 2011, Section 4.3).

1.2.2 The Mechanistic Account

On the grounds of these and other shortcomings of Salmon's approach, proponents of the CM model of explanation have developed alternative accounts. In what follows I consider one of these accounts, namely the *mechanistic conception* of explanation put forward by Craver (2006, 2007a, 2008, forthcoming a), Bechtel (e.g. 2006, 2008, 2009, Bechtel/ Abrahamsen 2005), Glennan (2002, 2010), and others.[22]

The first thing to note is that there exists an important difference between Salmon's CM model and the mechanistic account. As mentioned before, Salmon recognizes two aspects of causal explanations, the etiological and the constitutive aspect. He argues that in many cases causal explanations possess both aspects (1984b, 297). However, the examples which Salmon discusses show that he is primarily concerned with *etiological* causal explanations, that is, with explanations that describe the antecedent causes that lead up to the phenomenon in question. This is intelligible because examples of this kind are the ones that occupy center stage in the general debate about causation. To mention a few of them: depressing the gas pedal on a car explains why it accelerates, Suzy's throwing a rock at a window and the rock hitting the window explains why it breaks, smoking explains the development of lung cancer; the poisoning rather than the shooting explains the death of a man, and so on.[23] By contrast, the mechanistic account focuses not on etiological, but on

---

[21] In a recent paper Salmon concedes this point (1997). He replies that the notion of causal and explanatory relevance can be captured by his CM model and information about statistical relevance relationships.

[22] This focus on mechanistic models of explanation is due to my interest in reductive explanations. It does not imply that other recent versions of the CM model are less convincing or less valuable.

[23] Granted, in most of these cases there are also explanations available that are not only etiological, but constitutive as well. For instance, the explanation of the death of a man by him being poisoned may involve a description of the causal mechanism of how the poison is absorbed, how the cellular metabolism is affected by the poison, and how this causes certain symptoms of poisoning. But this is not how the explanans in these cases typically is characterized. Usually, philosophers discuss whether the poisoning causes (and causally explains) the death of a man, not whether certain underlying physiological mechanisms cause (and causally explain) the death of a man. To put it in a nutshell, the way that Salmon and others *in fact* discuss these

constitutive causal explanations, that is, on explanations that describe the underlying causal processes and interactions that constitute the phenomenon to be explained. This is due to the fact that mechanistic explanations are regarded as being constitutive, not etiological causal explanations. For instance, Craver states:

> Mechanistic explanations are *constitutive* or componential explanations: they explain the behavior of the mechanism as a whole in terms of the organized activities and interactions of its *components*. (2007a, 128)[24]

Proponents of the mechanistic conception of explanation state that "to give a description of a mechanism for a phenomenon is to explain that phenomenon, i.e., to explain how it was produced" (Machamer, et al. 2000, 3). Although one sometimes gets the impression that the mechanists assume an almost universal range of application of their account,[25] on closer inspection it becomes apparent that nobody really advocates this thesis in its radical version.[26] Most mechanists restrict their account to a certain scientific field (e.g. neuroscience, cell biology, etc.), and allow other kinds of explanations besides constitutive mechanistic explanations (e.g. mathematical explanations, topological explanations, "aggregate explanations", "morphological explanations" Craver 2007a, 162). Hence, by restricting the scope of their account the mechanists avoid a standard criticism that is raised against CM models, namely that there are *non-causal* explanations (and, thus, also non-mechanistic explanations) in the sciences, too.

Let us come back to the notion of a mechanistic explanation. According to Craver, the behavior of a mechanism as a whole (i.e. the explanandum phenomenon) is explained by describing how certain entities and activities are organized together such that they jointly produce the behavior to be explained (2007a, 2-9, 121-162). This general characterization of mechanistic explanations can be specified in the following way: first, mechanistic explanations are a special kind of *part-whole* or constitutive explanations. That is, the behavior of a mechanism as a whole is explained in terms of its parts (i.e. entities) and the activities they are engaged in. Second, in mechanistic explanations a certain behavior of a

---

examples shows that they focus on etiological causal explanations. But this does not imply that the phenomena they discuss could, in principle, not be explained in an etiological-constitutive manner, too.

[24] Craver stresses this at several places in his book. Against this background it is somewhat irritating that he also claims that "[s]ome mechanistic explanations are etiological, […] [o]ther mechanistic explanations are constitutive" (2007a, 107f). Craver seems to use two notions of mechanistic explanations: mechanistic explanations in the *narrow sense* are identical with constitutive causal explanations (and this is the kind of mechanistic explanations he develops his account for), whereas mechanistic explanations in the *wide sense* can encompass other types of causal explanations as well (which is why Craver can argue that functional explanations are mechanistic as well; forthcoming b).

[25] Several mechanists try to broaden the scope of the mechanistic conception. For example, Craver argues that etiological and functional explanations are subtypes of mechanistic explanation (2007a, 107; forthcoming b), Skipper and Millstein (2005) apply the mechanistic account to natural selection explanations, and Glennan (2010) claims that it holds for historical explanations, too.

[26] For instance, Craver states: "Constitutive explanation is one important kind of explanation in neuroscience. But saying so does not commit me to the view that all explanations are constitutive." (2007a, 162)

system is explained by reference only to those parts of the system that work together to bring about this behavior. For instance, protein synthesis is explained by reference to DNA strands, ribosomes, and amino acids, but not by appealing to sarcoplasmic reticula, centrioles, or the cytoskeleton. Craver introduces the notion of "*constitutive relevance*" (2007a, 139; my emphasis; see also 2007b) to distinguish those parts of the system (together with the activities they perform), which are parts of the mechanism in question (and which are, thus, constitutively relevant to the behavior of the mechanism to be explained), from other, explanatorily irrelevant factors. Third, the entities and activities that make up a mechanism are spatially and temporally *organized* in a specific way (Craver 2007a, 134-139). This is why mechanisms differ from mere aggregates (that are literally the sum of their parts), from mere spatial arrangements, and from mere temporal sequences. Spatial organization means, for example, that the entities of a mechanism are localized in certain areas of the system, that they exhibit specific sizes, shapes, and spatial orientations towards each other. 'Temporal organization' refers the fact that the causal process that brings about the behavior of the mechanism can be subdivided into certain stages with a particular order, rates, and durations. Each of these stages involves specific activities and interactions between certain entities. Together the spatial and temporal organization sustain the "active organization" (Craver 2007a, 137) of the mechanism's components, that is, their working together to bring about a particular behavior. The description of how the components of a mechanism are actively, spatially, and temporally organized constitutes the major part of a mechanistic explanation.

Finally, let me add some critical remarks on the mechanistic account of explanation. On the one hand, one can accept the general framework of this account and challenge its details. For instance, one might argue that the concept of a mechanism is still not sufficiently clarified, that it rests on problematic concepts and metaphysical assumptions (like the mechanists' notion of productivity and their thesis that activities are distinct ontological kinds, e.g., special types of causes), that it often remains unclear on which concept of causation the mechanistic account is based, that Craver's notion of constitutive relevance fails to offer a convincing account of what it means for a factor to be explanatorily relevant to a certain explanandum phenomenon, and so on.[27] On the other hand, one can question the endeavor of the mechanists as a whole. If they, as it seems, restrict their account to certain kinds of explanations (i.e. to constitutive causal explanations) it is questionable whether they offer a real competitor to the CL model. It seems to me that a real competitor must share the universalistic aspiration of the CL model, that is, it must claim to hold for scientific explanations in general, not only for constitutive causal explanations in a certain scientific field.[28] On the contrary, if the mechanistic model is thought to be a universal model

---

[27] I won't take a stance on these points here since this would prolong this overview unnecessarily. For further details about my own view see Chapter II, Section 3.2.2 and Chapter V, Section 1.2.2.

[28] According to this reading, a pluralistic account of explanation (like the one I develop in Section 2.2) is universalistic in aspiration, too.

of explanation that applies to *all* scientific explanations it is vulnerable to counterexamples since it does not capture etiological causal explanations.[29] In addition, it is faced with the general objection to CM models that they do not account for non-causal explanations.

1.2.3 Concluding Remarks

All in all, proponents of a CM model agree that the nature of (scientific) explanation is causal. They reject the claim that explanations are arguments and that the explanatory force in explanations arises from rendering the explanandum phenomenon nomically expectable, or from subsuming it under a general law. Rather, what renders a description explanatory is the fact that the causes, which lead to or make up the explanandum phenomenon, are adequately represented. In other words, the basic idea behind the CM model is that to explain a phenomenon amounts to situating it within the causal structure of the world.

One should be very cautious not to interpret the opposition CL model versus CM model as an opposition between explanation *by* laws/generalizations and explanation *without* laws/generalizations. This is not what marks the difference between these two accounts. The CM model does not exclude generalizations (which might count as laws or not) from occupying center stage in explanations. As a matter of fact, most causal explanations in science, in fact, contain generalizations.[30] The difference is that the CM model allows generalizations to figure as parts of the explanantia (provided that they are causal), but does *not require* them, whereas according to the CL model laws (plus deductive relation between explanans and explanandum) are *necessary* for explanation.

Besides the specific criticism that can be raised against particular accounts, CM models in general encounter three objections: first, they are opposed for relying on the concept of causation that has not been sufficiently specified so far; second, it is argued that the CM model provides no account of what it means for a causal factor to be explanatorily relevant; third, the requirement that explanations must embed a phenomenon into the causal structure of the world is too strong since there are also non-causal explanations in science. We have seen that Salmon and the mechanists try to meet the first objection by relying on elaborated theories of causation (i.e. Dowe's and Woodward's). Furthermore, I outlined that Salmon's CM model in fact fails to yield a notion of explanatory relevance. As opposed to this, Craver offers a promising notion of explanatory relevance (interpreted as "constitutive relevance" 2007b), although it is confined to the class of constitutive causal explanations. Finally, non-causal explanations remain forceful counterexamples, which, at least, restrict the scope of CM models.

---

[29] A possible way out would be to integrate the mechanistic account into a general causal conception of explanation like Woodward's (2003) or Streven's (2008).

[30] In line with this, see for instance the claim by Machamer, et al. that most mechanisms produce a particular behavior in a regular way, and that mechanistic explanations must describe these regularities (2000, 3; see also Darden 2008, 964f).

1.3 What Makes an Account of Explanation "Ontic"?

In this section I consider in more detail a specific characteristic that is ascribed to CM models. Most proponents of a CM model of explanation follow Salmon and emphasize the *ontic* character of their account (e.g. Craver 2007a, 27, 200; forthcoming a; Glennan 2002, 343; Strevens 2008, 6f, 43).[31] But what exactly does the term 'ontic' mean? That is, what makes a model of explanation ontic? Let us start by examining how Salmon employs this concept.

Salmon distinguishes three basic conceptions of scientific explanation:[32] first, *epistemic* conceptions, which treat explanations as arguments that render the event-to-be-explained expectable in virtue of the explanatory facts,[33] second, *modal* conceptions, which hold that there exists a relation of nomological necessity between the antecedent conditions and the event-to-be-explained, third, *ontic* conceptions, according to which an event is explained by exhibiting it as occupying its (nomologically necessary) place in the discernible patterns of the world.[34] Those philosophers who adopt the ontic conception (early proponents were Salmon, Railton 1981, and Coffa 1974) generally regard the pattern into which events are to be fitted in as a causal pattern. Thus, Salmon states that "[t]o give scientific explanations is to show how events… fit into the causal structure of the world" (1977, 162). More than a decade later Salmon summarizes his position as follows:

> According to the ontic conception, the events we attempt to explain occur in a world full of regularities that are causal or lawful or both. These regularities may be deterministic or irreducible statistical. In any case, the explanation of events consists in *fitting them into the patterns that exist in the objective world*. (1989, 121f; my emphasis)

I suggest that there are two different readings of Salmon's notion of an ontic conception of explanation. First, one can argue that what makes an account of explanation ontic is the thesis that explanations are *objects or facts* that exist *in the world*, independently from scientists discovering them, from inquirers requesting them, or from speakers uttering them. In the framework of the CM model these explanatory facts in the world are assumed to be causes or parts of the causal structure of the world. This interpretation of 'ontic' can most clearly be found in the work of Craver (2007a, 27, 33, 200; forthcoming a) and Strevens (2008, 6f, 43). For instance, Strevens states that explanations are "something out in the world, a set of [causal] facts to be discovered" (2008, 6). And Craver makes it even more explicit:

---

[31] Strevens deviates from Salmon's terminology and calls his account an "ontology-first approach to explanation" (2008, 7).

[32] These three conceptions of explanation represent different ways to interpret Laplace's view of explanation (1951, 4).

[33] Salmon characterizes inferential conceptions, like Hempel's DN and his IS model, as well as erotetic conceptions, like van Fraassen's pragmatic account of explanation (which will be discussed in Section 2.2.2), as epistemic conceptions of explanation.

[34] For further details on these three conceptions see Salmon 1984a, 15-20, 84-134.

> [T]he term explanation refers to an objective portion of the causal structure of the world, to the set of factors that bring about or sustain a phenomenon (call them objective explanations). […] Objective explanations are not texts; they are full-bodied things. They are facts, not representations. They are the kinds of things that are discovered and described. There is no question of objective explanations being 'right' or 'wrong', or 'good' or 'bad'. They just are. (2007a, 27)

But even Craver and Strevens do not deny that explanations sometimes are epistemic units, like descriptions, representations, models, or explanatory texts that are used to convey information or communicate scientific knowledge.[35] However, they emphasize that one should give precedence to the ontological sense of explanation. They claim that it is important to shift the attention away from the representations used in explanations and toward the causal structure of the world since only "objective explanations" provide "systematic clues about the nature of explanation itself" (Stevens 2008, 6) and, thus, constitute "the correct starting point" (Craver 2007a, 27) for developing an account of explanation.

Second, according to a less radical reading of the ontic conception, explanations are not mind-independent things in the world itself, but exclusively epistemic units (i.e. descriptions, representations, explanatory texts). That is, they are not the causes or parts of the causal structure of the world themselves, but rather descriptions or representations of these causes and partitions of the causal structure. This is the reading Glennan endorses:

> Causal-mechanical explanation exemplifies what Salmon calls the ontic conception of explanation. Explanations are not arguments, but are rather *descriptions* of features of a mind-independent reality – the causal structure of the world. (2002, 343; my emphasis)

Similarly, Bechtel and Abrahamsen argue:

> Salmon identifies his approach to explanation as ontic insofar as it appeals to the actual mechanism in nature… Salmon's insight is important […]. But it is crucial to note that offering an explanation is still an *epistemic activity* and that the mechanism in nature does not directly perform the explanatory work. (2005, 424f; my emphasis)[36]

But if it is not the thesis that explanations are things in the world, what is it then that makes an account of explanation ontic? The quotation from Glennan gives a hint. He states that according to the ontic conception explanations are not arguments, but descriptions of the causal structure of the world. That is, the explanatory force arises not from logical features of statements or from relations of nomic expectability. Instead, what makes a description explanatory is that it truly represents certain features of a mind-independent reality, namely the causal structure of the world. In sum, according to this reading, the characterization of a conception of explanation as ontic, rather than epistemic, does not result from the

---

[35] But note that this does not commit one to the claim that explanations are "communicative acts", as Strevens alleges (2008, 6f).

[36] A similar statement can be found in Bechtel 2006, 31f.

ontological status that is ascribed to explanations (i.e. that they are facts in the world rather than representations of these facts). Rather, the label 'ontic' indicates a certain thesis about the *determinants of the explanatory force* of an explanation. Proponents of an ontic conception claim that what determines whether a description counts as an explanation or not is whether it truly represents certain facts in the world. They deny that the explanatory force stems from certain logical relations that hold between the statements by which the explanation is expressed. In short, ontic accounts trace the difference between explanations and non-explanations, ultimately, back to *differences in the world*, and not to logical features of linguistic entities. For instance, the description of a sarcoplasmic reticulum releasing calcium ions when the membrane is polarized does not explain how a cell synthesizes proteins because the sarcoplasmic reticulum in the world does *not causally affect* protein synthesis. The failure of explanation is not due to the fact that the statement that describes protein synthesis cannot be logically derived from the generalization that describes the behavior of sarcoplasmic reticula.

My opinion is that the ontic character of an account of explanation should be understood in accordance with the second, less radical interpretation. The thesis that explanations are out there in the world, that they are identical with certain causes or causal structures, and that they exist independently of anybody asking questions and identifying phenomena to be explained, is misleading in several respects.

First, this sense of explanation is completely decoupled from understanding. Whether the tracing of the causes that lead to or make up a certain phenomenon renders this phenomenon intelligible and promotes its understanding is simply irrelevant for whether this part of the causal structure of the world is an explanation (for a certain phenomenon) or not. This contradicts our common sense notion of explanation, according to which the process of explaining a phenomenon is closely linked to the process of understanding why this phenomenon occurs.

Second, even proponents of the first reading, like Craver and Strevens, concede that the term 'explanation' has two different meanings. It refers not only to things in the world, but also to representations or explanatory texts. In my view, this preserves their position from being completely implausible. However, their claim that in order to grasp the nature of scientific explanation one must focus exclusively on explanations as things in the world is unmotivated and unconvincing. Even if one accepts their thesis that explanations sometimes are things in the world (which I reject) this does not imply that they are the *only* adequate target of philosophical analysis. Neither Craver nor Strevens offer a real argument for why this should be the case.[37]

---

[37] For example, Strevens simply states that "[e]ither sense may be given precedence" and that he follows "the lead of most philosophers of explanation" (2008, 6) in giving the ontological sense of explanation precedence. I doubt that Strevens is right in his assessment of what most philosophers of explanation do (see, e.g., Mitchell 2009; Woodward 2003, 2011; Brigandt forthcoming; etc.).

Third, Craver's way of speaking about "objective explanations" (2007a, 27) suggests that the difference between explanations as things in the world and explanations as representations is a difference between objective and subjective explanations. But this is simply not the case. Even if one adopts the second reading and stresses that explanations are representations or texts (as I do), one is not committed to the thesis that the explanatory force of a description is dependent on subjective factors. That is, whether or not a certain description of causes is explanatory or not can still be an *objective* matter.

In sum, I understand the notion of an "ontic" conception of explanation as indicating a certain thesis about the determinants of the explanatory force of an explanation. According to an ontic account, explanations are representations. But what determines whether a representation counts as an explanation or not is whether it truly represents certain *causal relations in the world*, not whether certain logical relations hold between explanans-statement and explanandum-statement. It might be that the more radical reading, according to which explanations are mind-independent things in the world, is an (at least partially) adequate interpretation of how Salmon employs the term 'ontic'.[38] But even if it adequately captured Salmon's position I would reject it on the grounds just described and propose the second reading as the way that *I* understand the notion of an ontic conception of explanation.

2 Explanation and Reduction

The goal of the last section was twofold. It served, first, to introduce the two main accounts that are discussed in the debate about explanation – namely the CL model and the CM model – and, second, to clarify in what sense CM models of explanation are referred to as 'ontic'. The latter issue is relevant because I characterize my own account of explanatory reduction in biology as ontic, too. In Chapter V, Section 6, when I present my account or reduction, I will pick up on this issue again, explicate why I label my account ontic, and elucidate in which respect my usage of the term 'ontic' is similar to the one that is present in the debate about explanation. The introductory notes on CL and CM models of explanation in turn provide the basis upon which I can now address more specific questions about explanation that concern the issue of reduction.

Recall the assertions from which we started. As we have seen, some authors claim that the debate about reduction is, basically, a debate about what constitutes a good or adequate explanation (van Regenmortel 2004a, 145). Others explicitly deny this. They state that questions about explanatory reduction cannot be decided by settling for a certain position in the debate about explanation (Rosenberg 2006, 41). So, which of them is right? I have already indicated that this needs to be judged *separately* with respect to reduction and with respect to reductionism. Roughly speaking, the question of what explanatory reduction

---

[38] Passages that support such an interpretation can be found, for example, in Salmon 1989, 86 and 1992, 35.

in biological practice is, that is, of what constitutes the reductive character of biological explanations, does *not* boil down to the question of what characterizes an adequate explanation. That is, what makes an explanation reductive is *different* from what makes a description explanatory – although there are interdependencies between how one answers the question of (explanatory) reduction and the question of explanation (see Section 2.1). In contrast, the question of whether explanatory reductionism is true or not in fact amounts to particular questions about explanations. In Section 2.2.1 I examine what these particular questions are. With regard to reductionism (but not with regard to reduction) I thus side with van Regenmortel and oppose Rosenberg. In what follows I first dwell on the relation between my account of explanatory reduction and questions about explanation (Section 2.1). Then I specify the questions about explanation that are relevant to discussions about explanatory reductionism (Section 2.2).[39]

2.1 Explanation and My Account of Explanatory Reduction

2.1.1 Different Questions

In Section 1 we became acquainted with two accounts of (scientific) explanation – the covering law (CL) model and the causal mechanical (CM) model. Both accounts differ in what they regard as the nature of explanation. In other words, they provide different answers to the question of what constitutes the explanatory force of explanations and distinguishes them from mere descriptions. Proponents of the CL model treat explanations as arguments and trace the explanatory force back to the deductive relation that exists between certain law statements, antecedent conditions and the explanandum phenomenon. According to them, the nature of explanation is nomic expectability (respectively pattern/law subsumption). Contrary to this, the CM model identifies causation as being crucial for explanation. Explanations are not taken to be arguments, but rather descriptions of the relevant causes (or as the causes itself; see Section 1.3) that lead to or make up the explanandum phenomenon. According to the CM model, to explain a phenomenon means to situate it within the causal structure of the world.

The point I want to stress here is that this question of explanation (i.e. is the nature of explanation nomic expectability or causation?) is *different* from the question of what makes an explanation reductive. The question of explanation serves to distinguish explanations from other kinds of scientific achievements (i.e. from mere descriptions or from purely predictive models). That is, accounts of explanation specify the conditions under which descriptions are explanatory, respectively, under which explanations are successful or adequate. By contrast, the question of explanatory reduction addresses a distinction *among* (adequate or successful) explanations, namely the distinction between reductive and non-reductive explanations. Consequently, answering the question of explanation does *not*

---

[39] As noted before, I focus on those issues that are relevant to my account of explanatory reduction.

answer the question of (explanatory) reduction. What we need in order to answer the question of reduction are *additional* criteria that an explanation must satisfy to count as a reductive explanation.[40] These additional criteria will be specified in the subsequent chapter. The difference between the two questions can be summarized as follows:

> *Question of Explanation:*
> According to which criteria is the set of adequate descriptions distinguished into the two subsets adequate explanations and non-explanations?
>
> *Question of (Explanatory) Reduction:*
> According to which criteria is the set of adequate explanations distinguished into the two subsets reductive explanations and non-reductive explanations?

However, even if the task of developing an account of explanatory reduction is different from the one of developing an account of explanation, one might wonder whether they are *completely independent* from each other. In particular, the question arises whether my account of explanatory reduction is compatible with any account of explanation. Having in mind my rejection of Nagel's model of theory reduction as being inadequate to biological practice (see Chapter II, Section 4) one might query whether my account is compatible with a CL model of explanation (since Nagel's approach is based on Hempel's DN model). I agree that my account of explanatory reduction goes much more smoothly with a CM model of explanation. My account of reduction and the CM model of explanation both treat formal issues, like deductive relations between statements, as being irrelevant for reduction/ explanation. Instead, they focus on substantive issues, like the causal relations that exist in the world (in the case of explanation) or the part-whole and system-environment relations that exist in the world (in the case of reduction). Moreover, I must admit that if I had to choose a side I would opt for the CM model because it encounters less devastating objections than the CL model (see Section 1.1.2), although it has its shortcomings, too (see Section 1.2.3). But I also think that it is, in principle, *possible* to combine my account of explanatory reduction with a CL model of explanation (if one tolerates some tensions). A proponent of the CL model could insist that the explanatory force of an explanation arises from the logical derivation of the explanandum phenomenon from certain law statements and statements of antecedent conditions. Nevertheless, he could (contrary to what Nagel claims) confess that formal relations do *not suffice* to distinguish reductive from non-reductive explanations and adopt my account of explanatory reduction. Such a combination of a formal account of explanation with a non-formal account of explanatory reduction gives rise to some tensions, but it would be feasible.

At this point one might criticize me for remaining as uncommitted as I do. Would it not strengthen my argumentation to take a clear position and more explicitly argue for the

---

[40] This point is also emphasized by Sarkar (1998, 9).

CM model of explanation? I think this is not the case. The question of what makes an explanation reductive is distinct from the question of what makes a description explanatory and the former can be answered independently from the latter. Since my aim is to develop an account of explanatory reduction that convinces as many people as possible, it is beneficial for me to remain *as neutral as possible* concerning the nature of explanation and not to side with one of the competing accounts of explanation.

Let me conclude with highlighting a respect in which this proclaimed neutrality of my account of reduction is constrained. In Section 1.3 I have argued for a certain interpretation of the notion of an ontic conception of explanation. In my view the thesis that explanations are the causes in the world itself and not merely representations or descriptions of these causes is deeply flawed. In this respect my account of explanatory reduction is not based on a completely neutral view of explanation (other constraints of this neutrality will be disclosed in Section 2.2). It takes for granted that explanations are epistemic units (which consist of two parts, namely the explanandum that is explained by the explanans), not things in the world. However, this decision excludes only more radical versions of the CM model (like Craver's and Strevens'). It leaves the question of explanation (i.e. is the nature of explanation nomic expectability or causation?) unanswered.

2.1.2 Some Matters of Terminology

As I have pointed out in the previous section, my account of explanatory reduction remains neutral with respect to what the nature of explanation is. I neither claim that explanations must refer to laws or generalizations, nor that all explanations must be causal. This is the reason why I use the, admittedly, unspecific notion of explanatorily relevant "*factors*"[41]. This terminology has two merits: first, it leaves open whether these factors are causal or not. Since most reductive explanations in the biological sciences exhibit a temporal character (Hüttemann/Love 2011; Love/Hüttemann 2011) in most cases these factors will, in fact, be causal factors, and the corresponding explanations will be reductive causal explanations. However, there may be reductive non-causal explanations in biology, too (e.g. gene-selectionist explanations[42]). My way of speaking about "factors" that are referred to in the explanans leaves room for this kind of reductive explanation.[43] Second, the notion of explanatorily relevant factors is uncommitted with regard to the question whether the explanans constitutes a generalization or contains law statements or not. Advocates of the

---

[41] I use the term 'factor' as an *ontological* notion, including causes, causal relations, causal processes, regularities, universals, or whatever one's ontology demands. That is, sets of factors are what is being described or represented in explanations, not what constitutes the explanation itself.

[42] That is, evolutionary explanations that explain evolutionary processes exclusively by appealing to selection processes at the level of genes (see Dawkins 1976; Wimsatt 1980). However, it is controversial whether natural selectionist explanations are to be characterized as non-causal explanations (e.g. Sarkar 2005, 117-143).

[43] Despite this, reductive *causal* explanations will occupy center stage in my analysis since they are the paradigmatic and most important examples of reductive explanations in biology. Non-causal explanations are more often examples of non-reductive explanations (e.g. the topological explanation of dominance; see also Sarkar 1998, 169-173; 2005, 98).

DN model will insist that only those factors count as explanatorily relevant that are referred to in the general law statement(s) and the statements of antecedent conditions from which the explanandum phenomenon can be deduced. Opposed to this, proponents of a CM model will object that there may be explanantia that contain no law statement or even no generalization at all. My usage of the term ‘factor’ is uncommitted to both.

It may strike the attentive reader that I sometimes deviate from speaking about “factors” that are described in the explanans and, thus, denoted as explanatorily relevant. In addition, I also state that a phenomenon or a behavior of a system is explained by appealing to the parts of the system and the interactions between them or by reference to certain entities, their properties, and the activities they are engaged in. Contrary to this, I refrain, for instance, from speaking about events or processes as represented in the explanantia.

$$\begin{tikzcd} \textbf{explanandum} \arrow[d, "\text{(represents)}"] & \textbf{explanans} \arrow[l, "\text{(explains)}"'] \arrow[d, "\text{(represents)}"] \\ \text{phenomenon or behavior of a system} & \text{(explanatorily relevant) factors or entities, their properties + interactions/activities} \end{tikzcd}$$

This terminology is common in the philosophy of biology, in particular, in debates about reduction and part-whole or mechanistic explanations. However, it is important to note that it might be replaceable because there is no strict ontological thesis standing behind it. To put it another way, by using this terminology I neither claim that, besides entities that change their properties, no real causal processes or “activities” (Machamer, et al. 2000, 4) exist,[44] nor do I reject (or advocate) the ontological thesis that events exist and that they are properties of a spatiotemporal regions (Lewis 1986). My aim is simply to remain ontologically neutral and to use such concepts in my analysis of reductive explanations that are common in the debate as well as descriptively adequate to biological practice (compare Chapter V, Section 1.2.2).

2.2 Explanation and Disputes about Reductionism

Let us now turn to the controversial issue whether the debate about explanatory reductionism amounts to a dispute about explanation, as van Regenmortel claims, or is distinct from general discussions about explanation, as Rosenberg asserts. In order to decide which of them has the better arguments we, at first, need to figure out what these disputes about explanation are, to which questions about reductionism shall amount to or shall be different from (Section 2.2.1). After having clarified this I argue that the dispute about

---

[44] I address the question whether the adequate ontological inventory of mechanisms are interacting entities (monism) or entities and activities (dualism) in Chapter V, Section 1.2.2. See also Machamer, et al. 2000; Tabery 2004; Torres 2008.

explanatory reductionism in fact boils down to answering specific questions about explanation, namely questions about the adequacy of higher- and lower-level explanations (Section 2.2.2). One of my main theses is that one's stance on this issue is mainly affected by the stance one takes on the pragmatics of explanation. Thus, in Section 2.2.3 I examine van Fraassen's pragmatic theory of explanation as well as the objections Rosenberg raises against erotetic conceptions of explanation. My goal is to reject Rosenberg's criticism by developing a refined version of van Fraassen's pragmatic account. This enables me also to give a negative answer to the question whether there is just one right level of explanation for each biological phenomenon (Section 2.2.4).

Before I begin, let me address a possible objection. Having in mind my emphasis that my goal is to develop an account of reduction, not to dispute the truth of reductionism (see Chapter II, Section 1), one might challenge the relevance of this entire section. Are questions about explanation, which concern the issue of reductionism, not completely irrelevant to my work? I think this is not the case. The issues I approach in this section are *relevant* to my work of seeking an adequate account of explanatory reduction, even though they are only indirectly relevant. Views of explanation like Rosenberg's must be shown to be inadequate since they would restrict the set of "empirical data", on which my analysis is built, in an inappropriate way. If Rosenberg were right only molecular explanations would count as reductive explanations. But the explanatory practice in biology shows a much larger variety of reductive explanations (see Chapter V), which needs to be taken into account. The revised version of van Fraassen's pragmatic account of explanation that I propose ensures that all different kinds of explanation that can be found in actual biological practice are considered. Thus, the subsequent discussion is indirectly relevant to my work since certain stances on explanation are incompatible with my account of explanatory reduction.

2.2.1 Which Questions about Explanation Are Relevant to Reductionism?

Consider the few remarks of van Regenmortel first. He states that

> [d]issatisfaction with reductionism… arises because scientists nowadays increasingly question the validity of reductive explanations for providing an understanding of what is causally relevant in bringing about biological phenomena. (2004a, 145)

This statement reveals that van Regenmortel is primarily concerned with *in-practice* reductionism (compare Chapter II, Section 1). In his view, the truth of explanatory reductionism depends on whether biologists in fact assess reductive explanations as adequate or not (not on whether it can be argued that reductive explanations in principle are adequate). Accordingly, the challenge to antireductionists is to specify the grounds on which reductive explanations of biological phenomena prove to be inadequate[45] (Byerly

---

[45] Strictly speaking, I should say "grounds on which reductive *representations* of biological phenomena prove to be *non-explanatory*" since I use 'explanation' as a success term (see fn. 2). However, such a way of speaking is less intelligible, which is why I accept the inaccuracy involved in using the term 'inadequate explanation'.

2003). Following this line of reasoning, the dispute about explanatory reductionism in fact boils down to a dispute about explanation, namely to the question of what determines the *adequacy* of an explanation or, in other words, which norms of explanation exist in biological practice, on the basis of which descriptions are judged as explanatory.

Rosenberg's Critique of Antireductionism

If we look at Rosenberg's argumentation, it becomes even clearer *which* disputes about explanation are meant. But in order to clarify this, we need to go into greater detail. In his newest book on "Darwinian Reductionism" (2006) Rosenberg takes up a discussion that has been prevalent since the rise of the reductionism debate in the philosophy of biology. It was initiated by Kitcher (1984) who challenged the thesis that molecular biology provides an explanation of Mendel's second law of independent assortment. Kitcher's argument can be traced back to Hilary Putnam's (1975, 295-298) famous discussion of the peg-hole example. According to Putnam, the phenomenon that a particular cubical peg passes through the square hole in a board, but not through the round hole, can be adequately explained by considerations from geometry (i.e. that the round hole is smaller than the peg and that the square hole is bigger than the cross-section of the peg). By contrast, Putnam regards a physical explanation of this phenomenon in terms of "ultimate constituents" (1975, 296), which describes the peg and the board as rigid lattices of atoms and appeals to laws of particle mechanics and quantum electrodynamics, as not being an explanation at all (or, at least, a "terrible explanation" 1975, 296). Likewise, Kitcher claims that the general phenomenon that genes on non-homologous chromosomes assort independently during meiosis is explained in cytological terms, but not by reference to molecules. He argues that a molecular description of "PS-processes" (i.e. pairing and separation processes; 1984, 349), which ensure independent assortment, "objectively fails to explain" (1984, 350). In Kitcher's view this failure of explanation arises from "the loss of understanding through immersion in detail" (1999a, 206) and from the failure to recognize from the molecular perspective what all PS-processes have in common. Putnam's and Kitcher's arguments have been critically discussed by various authors and from different perspectives (e.g. reductionist as well as antireductionist).[46] What is common to all these discussions is that they concern the question of what the *right level of explanation* is. An explanatory antireductionist (like Kitcher) argues that a higher-level phenomenon (like the cubical peg passing through the square hole or the independent assortment of genes) can only be adequately explained on a higher level and that therefore these higher levels of explanation are autonomous towards lower levels. An explanatory reductionist (like Rosenberg) rejects this view and claims that even higher-level phenomena are adequately explained exclusively on lower levels, respectively on the lowest level (i.e. either the molecular or the physical level).

---

[46] To name only some philosophers of biology: Waters 1990, 131-134; 2008, 244-249; Schaffner 1993, 478-481; Sober 1999; Rosenberg 2006, 32-47.

Let us track Rosenberg's line of reasoning a bit further. He criticizes Kitcher's argumentation by imputing a particular view of explanation to him, which he calls the pragmatic, "erotetic", or "Protagorean" (2006, 36f) account of explanation. Rosenberg contrasts this erotetic view with other non-erotetic accounts of explanation, like the DN model, the unificationist approach, and causal models of explanations (2006, 36). He gives the following characterization of the erotetic view:

> the erotetic account of explanations treats them [i.e. explanations] as answers to 'why' questions…, which are adequate – that is, explanatory – to the degree they are appropriate to the background information of those who pose the why question and to the degree that the putative explanation excludes competing occurrences or states of affairs from obtaining. (2006, 36)

According to Rosenberg, this view of explanation is deeply flawed, which he regards as the reason why Kitcher's argumentation in favor of the autonomy of higher levels of explanation goes wrong. Rosenberg argues that the erotetic account of explanation "relegates antireductionism to the status of a claim about biologists, not about biology" (2006, 36). That is, it turns the question whether a description is explanatory or not into a "subjective" matter of whether the information that is included in the description is relevant to the inquirers' interests (2006, 35, fn. 3). For Rosenberg, this is a misleading picture of explanation. He counters that "there is such a thing like a complete and correct explanation *independent* of contexts of inquirers' questions" (2006, 44; my emphasis). In other words, he claims that the adequacy of an explanation is a completely "objective" matter since it is independent from contexts and interests. His quite radical thesis is that, at least in the "context of advanced biological inquiry" (2006, 47), only those explanations are adequate (i.e. explanatory), which are *maximally complete* at the *molecular* level. To come back to Kitcher's example: in Rosenberg's view the cytological explanation of Mendel's law of independent assortment is inadequate because of its "silence regarding crucial links in the causal chains to which it adverts" (2006, 45). However, it can be corrected and made adequate by filling in the links in the causal chain which information about molecules and their interactions, which is to say, by providing a molecular explanation. This gives rise to Rosenberg's general thesis that functional biology's ultimate how-possible explanations must be reduced to (and, in so doing, made adequate by) molecular proximate why-necessary explanations (for further details see Chapter III, Section 1.1).

Relevant Questions about Explanation

As I will elaborate below (see Section 2.2.2), the picture that Rosenberg draws is deficient in several respects. For instance, it is not true that pragmatic accounts of explanation turn the question of explanation into a mere "subjective" matter. Moreover, the pragmatic account does not stand in contrast to other accounts of explanation, like CL and CM models, but rather is combinable with them (as already Salmon has argued; 1989, 135-146). But let us put these issues aside for a moment. What matters to my concerns at this point, is that

Rosenberg's argumentation yields an answer to the question from which we started. Namely, what are these disputes about explanation, to which questions about reductionism allegedly amount to or are different from? The answer to this question is threefold. The first point displays which general kind of question is relevant to reductionism, the second point specifies this kind of question, and the third point reveals the most important determinant of why this kind of question is answered differently.

First, the discussions about explanation that matter to the issue of reductionism concern the *conditions of adequacy*, respectively the norms of explanations. That is, the question is whether specific kinds of representations of biological phenomena are assessed as explanatory in biological practice or not. Formulated in terms of in-principle reductionism this question would be whether it can be convincingly argued that these specific kinds of explanation are the only ones that are adequate/explanatory in principle. In other words, questions about reductionism bear on the central question of how to distinguish explanations from non-explanations (i.e. from inadequate explanations).

Second, as indicated by the term 'specific', not all discussions about the adequacy of explanations are relevant to the dispute about reductionism. Only those that concern *levels of explanation* are crucial. That is, the question about explanation that matters to reductionism is whether certain kinds of explanation, namely those whose explanantia are restricted to certain levels (i.e. higher-level and lower-level explanations), are adequate or not. For example, the question is whether the phenomenon that a particular cubical peg passes through a square hole in a board can only be adequately explained by considerations from geometry (i.e. by a higher-level description), or whether exclusively/also a representation of the interactions between the atoms or molecules that constitute the peg and the board (i.e. a lower-level description) is explanatory. Generally speaking, the question whether explanatory reductionism is true depends on whether certain levels of organization can be said to be *explanatorily prior* to others. Antireductionists emphasize the explanatory autonomy of higher levels of explanation either by arguing for the explanatory priority of these higher levels (as Kitcher) or by endorsing the equal explanatory status of different levels of explanation. In contrast, reductionists typically defend the explanatory priority of lower levels or the lowest level of explanation.[47]

Third, Rosenberg's argumentation furthermore reveals that there is a second major dispute in the debate about explanation, besides the discussion about whether the nature of explanation is nomic expectability (as the CL model assumes) or causation (as CM models presuppose). This second major dispute concerns the role that *pragmatic factors*, like explanatory interests, background information, etc., might play in determining what counts as an adequate explanation and what not. I think it is crucial to note that the debate about the pragmatics of explanation is much *more relevant* to the topic of reductionism than the

---

[47] Note that this dispute does not concern the difference between better and worse explanations, but rather the question of whether higher- and lower-level explanations are adequate (i.e. explanatory) at all.

debate about whether CL or CM models correctly capture the nature of explanation. Rosenberg's as well as Kitcher's argumentation clearly shows why this is the case. Their central concern is whether a particular higher-level phenomenon can adequately be explained on a higher level or on a lower level. How one answers this question is affected by one's stance on the pragmatics of explanation. That is, whether one advocates an erotetic account of explanation and treats pragmatic factors as (co-)determining the adequacy of an explanation or whether one rejects such a view. By contrast, how one assesses the explanatory force of higher and lower level descriptions, is largely unaffected by whether one regards nomic expectability or causation as more important to explanation.[48] Neither the DN model, nor the causal model of explanation as such, contains assumptions about the allegedly "right" level of explanation. Hence, if the debate about explanatory reductionism amounted to discussions about explanation it would amount to discussions about the pragmatics of explanation.

However, there is an important restriction of this third thesis, which needs to be stressed. Several proponents of a CM model of explanation include additional assumptions in their account (e.g. about the nature of causation or about the adequacy of explanations), which eliminate the neutrality of their account with respect to questions about levels of explanation. First and foremost, the *mechanistic account* of explanation (see Section 1.2.2) regards mechanistic explanations as constitutive causal explanations, in which a phenomenon is explained by appealing to the underlying causal processes and interactions that constitute the phenomenon. To put it in another way, in mechanistic explanations the behavior of a mechanism is explained in terms of its parts, which can be said to be located on a *lower level* than the phenomenon to be explained. Thus, the mechanistic conception takes into account only lower-level explanations.[49] Most mechanists confess that their account is limited in scope (i.e. that it does not capture all kinds of explanation). Thus, they do not deny the existence of non-lower-level explanations. One can, however, also observe more universalistic aspirations in the mechanistic literature. The thesis that *all* (scientific) explanations are or should be mechanistic explanations (and, thus, lower-level explanations) would clearly violate the neutrality of the CM model with respect to levels of explanation.

Let me invoke one more example. Strevens' version of the CM model, which he calls "Kairetic Account of Explanation" (2008, 117), contains the thesis that only causal models that are as deep as possible are "standalone" (i.e. adequate) explanations. A causal model is "deepened" by offering "a lower-level account of some causal process represented by the model" (2008, 129). According to Strevens, this deepening is "compulsory" (2008, 133)

---

[48] This is not to say that the question whether subsuming phenomena under laws or capturing causal relations is more important for explanation is irrelevant for the issue of reduction. Because Nagel's model of theory reduction presupposes the DN model, the adequacy of DN models is crucial for discussions about reduction that remain inside the Nagelian framework. However, if one abandons Nagel's understanding of reduction, the dispute between CL and CM models of explanation loses significance, too.

[49] This is not to say that all mechanistic explanations must be fundamental-level explanations, like molecular explanations or even physical explanations. For further details on this difference see Chapter V, Section 2.3.1.

because "in order to understand a phenomenon fully, you must grasp the workings of the relevant causal mechanism in fundamental physical terms" (2008, 130f). This is an obvious violation of the neutrality of CM models with respect to levels of explanation since Strevens claims that to adequately explain a phenomenon requires uncovering the causal relations on the lowest, physical level. However, CM models of explanation need not include these additional assumptions about the explanatory priority of certain levels (as, for example, Woodward 2003 shows). This is why the debate about whether the nature of explanation is nomic expectability or causation is not, by itself, relevant to disputes about the correctness of explanatory reductionism.

In sum, the disputes about explanation which are relevant to reductionism concern questions about the adequacy of explanations (i.e. which conditions need to be satisfied in order that a description counts as an adequate explanation). In particular, questions about the adequacy of higher- and lower-level explanations (i.e. whether certain levels are explanatorily prior to others) are the ones that are crucial. The answers one gives to these questions are affected by one's stance on the pragmatics of explanation. Accordingly, discussions about whether pragmatic factors influence the adequacy of explanations are highly relevant to the issue of reductionism. For this reason, this topic will be examined in more detail in Section 2.2.3.

2.2.2 Why the Dispute about Reductionism Amounts to a Dispute about Explanation

The results of the preceding section strongly suggest that the truth of explanatory reductionism depends on whether one can convincingly argue for the explanatory priority of lower levels or the lowest level of organization. Since this typically involves taking a stance on the pragmatics of explanation, discussions about the truth of explanatory reductions amount to discussions about whether erotetic or non-erotetic conceptions of explanations are more plausible. The way Rosenberg himself discusses these issues (see 2006, 32-55, 178-181) strongly supports this view.

However, in a footnote Rosenberg maintains the contrary. He explicitly claims that "the debate [about reductionism] cannot be a dispute about 'explanation', for example a disagreement about pragmatic, erotetic, Protagorean versus nonerotetic accounts of explanation" (2006, 41, fn. 6). The reasons Rosenberg adduces for this thesis are not very persuasive. He argues that the dispute about the pragmatics of explanation is a "general problem in the philosophy of science, not a problem about reductionism in the philosophy of biology" (2006, 41, fn. 6) and that discussions about reductionism should be kept apart from the further issue of explanation. According to Rosenberg, one way to do this is to borrow Railton's notion of an ideal explanatory text (1981) and to characterize antireductionism as the claim that such an ideal text need not include descriptions of molecular factors. Correspondingly, a reductionist like Rosenberg insists that the ideal text must advert to descriptions and generalizations of molecular processes. It seems to me that this is an incorrect characterization of the point on which antireductionists and reductionists deviate from each other. Contrary to what Rosenberg states an antireductionist can agree that the ideal explanatory text entails molecular descriptions. However, the antireductionist will emphasize that explanations can be fully adequate even if they encompass only those parts of the ideal text that are non-molecular (as higher-level explanations do). Antireductionists à la Kitcher will add that descriptions, which include only the molecular parts of the ideal text, are not explanatory at all. However, an antireductionist is not confined to this radical thesis. He can also argue that both higher- and lower-level explanations are adequate (with respect to different questions, in different contexts; see Section 2.2.3). By contrast, a reductionist will reject this thesis and claim that only the molecular parts of the ideal explanatory text constitute adequate explanations.[50]

But even if Rosenberg's characterization of how the dispute about explanatory (anti)reductionism can be recast in Railton's terms were adequate this would not change the fact that his arguments for why the dispute about reductionism does not amount to a dispute about the pragmatics of explanation are sketchy and unconvincing. He simply states that the question of explanatory reductionism is different from the question of explanation (which I agree on) and insists that they need to be kept apart. However, in the previous section (see Section 2.2.1) I have shown that the two questions cannot be kept apart and that even Rosenberg himself fails to do this. Let us now turn to the pragmatics of explanations and assess the arguments that can be offered in favor of a non-erotetic conception of explanation.

2.2.3 Pragmatic Dimensions of Explanation

Van Fraassen's Pragmatic Theory of Explanation

The most prominent pragmatic account of explanation was developed by Bas van Fraassen (1977, 1980).[51] Central to his approach is the thesis that the explanation of a phenomenon P is an answer to a why-question *Q*. The general form of the explanation-seeking question is: "Why (is it the case that) *P*?" (1980, 126) According to van Fraassen, why-questions are essentially contrastive. That is, the same question-statement can express different propositions depending on what P is contrasted with. For instance, we ask 'Why is the muscle fiber contracted (rather than relaxed)?' or 'Why is the muscle fiber contracted (rather than the spinal nerve)?'. Both of these questions have different answers because an answer to a why-question must adduce information that favors P_k (i.e. the *topic* of the question) in contrast to other members of the *contrast class* X = $\{P_1, \ldots, P_k, \ldots\}$. Furthermore, van Fraassen argues that each why-question is characterized by a certain "respect-in-which"

---

[50] Accordingly, both the antireductionist and the reductionist deviate from Raiton's view. Even a moderate antireductionist who admits that higher- and lower-level explanations are adequate does not agree with Railton on the point that an explanation is better the more parts of the ideal text it includes.

[51] Another important representative of the erotetic version of the epistemic conception of explanation (to adopt Salmon's terminology) is Peter Achinstein (1983).

(1980, 142) an answer to the question is requested. This *relevance relation R* determines what counts as a possible explanatory factor (i.e. as part of the answer *A* to the question *Q*). If *R* holds between the answer *A* and the couple $\langle P_k, X\rangle$, *A* is called relevant to *Q*. In sum, van Fraassen identifies any why-question *Q* as an ordered triple $\langle P_k, X, R\rangle$.[52]

What is peculiar to van Fraassen's account is that he regards the contrast class *X* and the relevance relation *R* as being dependent on the *context*, in which the question is being asked (1980, 130). That is, the proposition a question-statement expresses can only be determined relative to a given context (which includes a body of background knowledge *K*). And since explanations are answers to questions, the adequacy of an explanation depends on the context, too. This is why van Fraassen concludes that explanation is a "three-term relation, between theory, fact, and context" and that "there can be no question at all of explanatory power as such" (1980, 156), that is, independently from the context in which a certain explanation-seeking question arises.

Let us come back to Rosenberg's argumentation. As we have seen, he contrasts the erotetic view of explanation (like van Fraassen's pragmatic theory) with other accounts of explanation (like the CL model and CM models) that pretend to having identified the nature of explanation. Since van Fraassen himself presents his account as an alternative theory of explanation (namely as a "theory of why-questions"; 1980, 134) this characterization is comprehensible. However, I think Kitcher and Salmon made an important and correct assessment when they argued that van Fraassen's treatment of the pragmatics of explanation "should be viewed as a *supplement*, rather than a rival, to the traditional approaches to explanation" (1987, 328; my emphasis). Their judgment is based on the claim that van Fraassen does not impose any restriction on the relevance relation R (at least not in his formal account). Besides his claim that the kind of explanatory relevance R that is requested is determined by the context in which a question is being asked, van Fraassen provides no further criteria for distinguishing genuine relevance relations from non-permissible ones. Consequently, van Fraassen offers neither a solution to the traditional dispute between proponents of CL and CM models (each of which spells out explanatory relevance differently), nor an alternative view of how the notion of explanatory relevance must be specified.

However, Rosenberg could easily grant this. Even if van Fraassen's view is to be characterized as a supplement, rather than a rival, to traditional CL and CM approaches, the core of Rosenberg's criticism against erotetic views of explanation remains applicable. His claim is that the erotetic view is fallacious because it converts the question whether a particular description is explanatory or not into a "subjective" matter of the explanatory interests, aims, background information, and cognitive capacities of biologists (2006, 13f, 35f, fn. 3, 44, 179f). Contrary to this, Rosenberg takes the distinction between explanations

---

[52] For more details on van Fraassen's theory of explanation see 1980, 134-157. For a summary of van Fraassen's account see Kitcher/Salmon 1987, 317-319.

and non-explanations and, thus, the relation of explanatory relevance to be an "objective" matter, which is unaffected by pragmatic factors. So, does Rosenberg have a point?

Evaluating Rosenberg's Criticism

First of all, the cogency of Rosenberg's criticism needs to be assessed differently with respect to different elements of van Fraassen's approach. It seems to me that there is a crucial difference between, on the one hand, van Fraassen's claim that the context affects the contrast class X and, on the other hand, his thesis that the context determines the relevance relation R. Briefly speaking, I see no reason why accepting the former implies that the question of whether a description is explanatory or not is affected by pragmatic factors. This is only true with regard to van Fraassen's thesis about R.

Consider the former claim first. According to van Fraassen, the topic P_k and the contrast class X specify the question that is being asked and, in so doing, affect what counts as an adequate answer to this question (i.e. what counts as explanatory). A different formulation would be to say that P_k and X specify the phenomenon to be explained. Van Fraassen states that what belongs to the contrast class X is determined by the context in which the question is being asked. In other words, which exact phenomenon is the explanandum phenomenon depends on pragmatic factors. I see no reason why this claim by itself turns the question of explanation (i.e. whether a certain description is explanatory or not) into a pragmatic or even a "subjective" matter. Even though the choice of the explananda phenomena (i.e. how the phenomena to be explained are specified) may be affected by the research interests and the background knowledge in a certain scientific field, this does not imply that the explanatory relevance of certain factors is determined by pragmatic factors, too. One could insist that, as soon as the phenomenon to be explained is fixed (by specifying P_k and X), the distinction between representations that explain this phenomenon and those that are non-explanatory is independent from pragmatic factors.

In my view, even proponents of a non-erotetic account of explanation could agree to van Fraassen's demand, that in order to give an explanation the topic P_k and the contrast class X of the question that is being answered by the explanation need to be specified. Probably they would not talk about questions being answered, but rather about phenomena that are being explained. But the core of van Fraassen's thesis would remain the same: the explanatory force of a description depends on how exactly the phenomenon to be explained is specified. Hence, I think Salmon is (partially) wrong when he states that

> although this issue [i.e. that explanations are answers to why questions] is crucial for an advocate of the erotetic conception, it has little – if any – genuine significance for the proponent of the ontic conception. According to this latter conception… the form of the question requesting them [i.e. explanations] is not very important. (1989, 138)

Even if one endorses a strong interpretation of the ontic conception, according to which explanations are "preexisting [causal] facts to be discovered" (Strevens 2006; see Section

1.3), one must consent to the thesis that an explanation is always an explanation *of* a specific phenomenon, and that, thus, the specification of a phenomenon affects what counts as an adequate explanation of this phenomenon and what does not. That is, one does not need to regard explanations as acts of communication or as answers to questions before one can accept the core of van Fraassen's thesis about P_k and *X*. In sum, Rosenberg's objection against this part of van Fraassen's account proves to be untenable.

Let us turn to van Fraassen's second thesis that the kind of relevance relation *R*, on the basis of which an answer *A* can be said to be relevant to a question *Q*, is determined by the context, too. Kitcher and Salmon convincingly argue that, unless van Fraassen imposes some conditions on relevance relations, his theory is committed to the result that "almost anything can explain almost anything" (1987, 322). In order to avoid this, he must show that explanatory relevance "is not completely determined by subjective factors…, but subject to some kinds of standards or criteria" (1987, 324). Similar to Rosenberg, Kitcher and Salmon point to the bugaboo of relativism or subjectivism, which a theory of explanation must avoid. The underlying assumption is that the question whether certain factors are explanatorily relevant to a particular phenomenon or not cannot be answered differently from subject to subject depending on individual background knowledge and preferences. This would render the question of the adequacy of an explanation into a highly subjective and almost arbitrary matter. I agree that this would be an inappropriate view of explanation.

However, denying this kind of subjectivism does not commit oneself to Rosenberg's thesis that explanatory relevance is *completely unaffected* by the context in which an explanation is developed (and even less to his claim that only descriptions that are maximally complete at the molecular level are explanatory). At this point Rosenberg performs a black and white thinking that is fallacious. There is no either pragmatics-then-subjectivity or no-pragmatics-then-objectivity choice. Instead, there are different "intermediate" ways to improve van Fraassen's account by adding a substantial notion of explanatory relevance, which avoid the reproach of subjectivism, but hold to the influence of pragmatic factors on what counts as explanatorily relevant. In what follows I give a brief outline of the way that appears to be most promising from my perspective.

Refining van Fraassen's Account: Types of Explanation and Explanatory Pluralism

Let us start by recalling some of my methodological preliminaries. In Chapter I, Section 1 I argued that it is important to my project to take serious the explanatory practice we find in the biological sciences. The reductive and non-reductive explanations that are actually developed in biological practice constitute a significant part of the empirical basis a philosopher must work with when he tries to understand what explanatory reduction in biology is. Against this background it becomes clear that I must adopt a similar stance with regard to the concept of explanation. That is, it would be peculiar if I claimed that concrete cases of adequate and inadequate explanations were not of particular interest for giving a philosophical account of what determines the adequacy of an explanation. One consequence of this methodological decision is that one must take seriously also the bewildering variety of *different types of explanation* that are available in scientific or even in biological practice. In accordance with what I said about balancing specificity against generality (see Chapter I, Section 4) I think a philosophical account of explanation must accomplish both; it must reveal commonalities across disciplines, but also enable us to see why explanatory practice varies as it does across different disciplines.

One possible way to account for the different types of explanation existent in biological practice is to abandon the thesis that there is just *one* genuine (or “objective” Salmon 1984a, 131) relevance relation *R* that applies to all types of explanation. Instead, one could argue that there exists an entire *set* of genuine relevance relations, each corresponding to a certain type (or more types) of explanation. For instance, in evolutionary biology we find many statistical explanations, in which an evolutionary outcome is proved to be more likely given the actual conditions (e.g. fitness differences) than given some alternative conditions (Sober 2008). In ecology, phenomena are often explained as being instances of quantitative generalizations, that is, by developing mathematical models that represent the dynamics of a biological system (e.g. how the density of a population changes during time; Sarkar 2009). In several other biological fields, such as molecular, cell, and developmental biology, causal explanations are widespread. Causal explanations can be further classified into subtypes, like constitutive/ mechanistic explanations and etiological explanations, the latter of which might be characterized as difference maker explanations (see Woodward 2003; Waters 2007). Depending on the concept of function one endorses, functional explanations, which occupy center stage in many biological fields, are either categorized as a subtype of causal explanations or as a distinct type (see McLaughlin 2001; Sarkar 2005). This list could go on with topological explanations, in which for example the phenomenon of dominance is explained by reference to topological properties of reaction networks (Huneman 2010; Sarkar 1998, 169-173), or with historical explanations (as certain kinds of evolutionary explanations), which might also be characterized as a special kind of etiological explanation. In sum, philosophy of biology is far from having established a complete list of types of biological explanations (see also Brigandt forthcoming). What is more, philosophers have only started to specify the notion of explanatory relevance with regard to these different types of explanation (see, for instance, Craver’s work on “constitutive relevance” 2007a, 139; 2007b). Accordingly, it is still an open question, which relevance relation *R* belongs to the set of *genuine* (or “objective” Salmon 1984a, 131) relevance relations.

However, in order to reject Rosenberg’s subjectivism-reproach we are not reliant on having identified and specified all relevance relations that count as genuine. All we need is the thesis that there exists such a distinction between genuine and non-acceptable relevance relations and that this distinction can be drawn independently from pragmatic factors, such as the background knowledge and interests of individual question-askers. According to this view, not the distinction between genuine and non-genuine relevance relations is affected by pragmatic factors, but rather the choice of a particular relevance relation among the members of the set of genuine relevance relation.[53] For instance, the explanatory interests that are accepted in physiology affect the choice that the phenomenon of photosynthesis is explained by describing the underlying cellular mechanism (i.e. how photons are absorbed by photosystems, how the flow of electrons through an electron transport chains creates a proton gradient, how this gradient is used to synthesize ATP, etc.), not (only) by appealing to the overall equation for the light-dependent reactions (i.e. $2\ H_2O + 2\ NADP^+ + 3\ ADP + 3\ P_i + \text{light} \rightarrow 2\ NADPH + 2\ H^+ + 3\ ATP + O_2$).

But even the choice of which of the genuine relevance relations applies to a certain explanandum phenomenon is not affected by *subjective* pragmatic factors (i.e. by the background knowledge and preferences of individual subjects). Rather, it depends on the research interests and theories that are commonly accepted in an entire research program, by a whole research group, or in a certain scientific field. Thus, even though the choice of relevance relations is affected by the context of inquiry, this does not imply that the choice varies from one individual researcher to another, or that it is arbitrary.

In addition, the relevance relations *R* themselves constitute intersubjective and often transdisciplinary *standards* for deciding which information must be included as relevant in the explanation of a particular phenomenon and which information is to be excluded as irrelevant. That is, only the choice of *R*, but not the content of *R* itself is determined by pragmatic factors. For instance, pragmatic factors may determine whether a phenomenon is adequately explained by showing that it is an instance of a (causal) generalization, by describing the causal mechanism that underlies the generalization, or by representing just certain parts of this mechanism, like the actual causal difference makers. However, it is not the case that these different relations of explanatory relevance are to be specified in different ways, depending on the context in which they are applied.[54] In sum, contrary to what Rosenberg claims, advocating this refined version of van Fraassen's account does not convert the adequacy of an explanation into a "subjective" matter or into a mere question about the interests of biologists.

The view of explanation I just presented is closely connected to what has been called "pluralism about the character of scientific explanation" (Brigandt forthcoming, Section 2) elsewhere. Three different theses of *explanatory pluralism* need to be distinguished. The refined version of Van Fraassen's account is committed only to the former two.

---

[53] The choice of *R* is sometimes entangled with the specification of the explanandum phenomenon. This is the reason why the choice of *R* often is indicated by the form of the question about the phenomenon that calls for the explanation. For instance, some philosophers have claimed that explanations of *how* something does what it does call for mechanistic explanations, whereas explanations of *what* something does call for explanations that include contextual factors (e.g. Duprè 1993, 106; 2009, 37).

[54] If this were the case we would say that these different specifications constitute distinct relevance relations.

First, the weakest kind of explanatory pluralism holds that two explanations of different phenomena can be of different kinds, that is, they assume different relevance relations between explanandum and explanans. To put it another way, this kind of pluralism recognizes that there exist different types of explanation, which involve different relevance relations.[55]

Second, according to a stronger form of pluralism about explanation, not only different phenomena, but *one and the same phenomenon* can be adequately explained differently by presupposing different relevance relations. Using a terminology that is different from van Fraassen's one could say that a particular phenomenon can be explained in more than one way by pursuing different explanatory aims, by approaching different explanatory projects, or by accepting different values/norms about explanation.

Third, some pluralists have argued that the arising different explanations of the same phenomenon are mutually *incompatible* and thus cannot be integrated into a single, complete representation of the world (e.g. Mitchell 2003, 2009; Kellert, et al. 2006; compare Chapter I, Section 4.3). Other philosophers reject this latter kind of explanatory pluralism. They insist that all adequate explanations of a certain phenomenon can be integrated into something like an "ideal explanatory text" (Railton 1981, 246). I must admit that the picture Railton paints appeals to me. And it could easily be aligned with the refined version of van Fraassen's view of explanation presented above. According to Railton, explanations provide explanatory information concerning why *P* (1981, 240), and as such they typically possess a partial character. That is, explanations single out or shed light on a particular part of some ideal explanatory text, which contains all information that is explanatorily relevant to the explanandum phenomenon *P*. In my words, explanations single out a certain part of the ideal text that is relevant to *P* due to a specific relevance relation *R*. In comparison, the ideal text contains all information that is relevant to *P* according to all genuine relevance relations *R* that are applicable to *P*. The only assumption one needs to abandon is Railton's thesis that explanations are better the more complete the part of the ideal text is that they constitute. However, despite the appeal Railton's picture clearly has, I also share the pluralist's skepticism of the assumption that all adequate explanations of a specific phenomenon can be integrated neatly into one coherent ideal explanatory text. If one considers the case studies presented by Kellert, Longino, and Waters (2006) the thought arises that Railton's picture might be too good to be true. But this is not something I need to decide in this book.

To conclude, van Fraassen's pragmatic theory of explanation can be refined in such a way that it avoids the objection of being an "'anything goes' account of explanation" (Kitcher/Salmon 1987, 328) and the criticism that it relegates the adequacy of an explanation to a "subjective" matter or to a mere question about the background knowledge

---

[55] This weak form of explanatory pluralism might even be compatible with the claim that, despite this diversity of explanations, there exists an underlying notion of explanatory relevance that is true for all (scientific) explanations.

and explanatory interests of biologists (Rosenberg 2006). However, explanation inevitably has pragmatic dimensions and this is an important fact to note. The research context determines which specific phenomenon *P* is to be explained (i.e. what the topic P_k and the contrast class *X* of the explanation-seeking question *Q* is) and it affects according to which genuine relevance relation *R* certain information are included in the explanans and others are excluded. But pragmatic factors neither influence which relevance relations at all count as genuine, nor do they influence how these relevance relations are to be specified. This opens up the possibility, on the one hand, to acknowledge that the explanatory relevance of a factor to a phenomenon is context-dependent – since the choice of the relevance relation is context-dependent – and, on the other hand, to insist that "the difference between a genuine explanation and a nonexplanation is perfectly objective" (Sober 1999, 550) – since the distinction between genuine and non-genuine relevance relations is independent from pragmatic factors. This leads us to the final question in what way this view of explanation affects questions about levels of explanation.

2.2.4 Is There a "Right" Level of Explanation?

In the previous section I emphasized the diversity of types of biological explanation, which I conceived as indicating a variety of different genuine relevance relations (e.g. statistical relevance, constitutive relevance, etiological relevance, mathematical derivability, fulfilling a function, etc.). Further research on this issue must show which candidates represent genuine relevance relations, how they are related to each other, and how they can be specified. Let us now narrow our view and focus on causal explanations.

The first thing to note is that *causal relevance* must be distinguished from *explanatory relevance*. The notion of causal relevance serves to distinguish what the causes of a particular phenomenon or event are and which factors are causally irrelevant to this phenomenon. Many different proposals have been made in order to specify what a cause is. However, these different theories of causation will not be investigated here. What is of concern to debates about explanatory reductionism is not how causes are distinguished from non-causes, but rather how a certain distinction among causes is drawn, namely how explanatorily relevant causal factors are differentiated from explanatorily irrelevant ones. In short, the picture that underlies discussions about reductionism in philosophy of biology is that there exist three types of factors:

(1) *causally irrelevant* factors,
(2) *causally relevant* factors that are explanatorily irrelevant, and
(3) causally and *explanatorily relevant* factors.

For instance, the working of cell organelles such as centrioles is said to be causally irrelevant for the synthesis of proteins in a cell (i.e. centrioles belong to (1)). By contrast, mitochondria are causally relevant to protein synthesis because they provide the energy (ATP) that is needed for several steps of protein synthesis (i.e. mitochondria belong to (2)). However, according to most notions of explanatory relevance mitochondria will not be regarded as explanatorily relevant. For example, Craver's notion of compositional relevance (2007a, 139-160; 2007b) treats only those factors as explanatorily relevant that contribute to producing the phenomenon to be explained (in other words, relevant factors and explanandum phenomenon need to be mutually manipulable). Accordingly, mitochondria are irrelevant because they make no direct contribution to bringing about the synthesis of proteins. The contrary is the case with regard to causal factors such as ribosomes, DNA and m-RNAs (which belong to (3)). They are both causally and explanatorily relevant.

According to this view, the natural world exhibits a particular causal structure, which can be discovered and represented. The role of the concept of *explanatory relevance* is now to specify which causal relationships in a phenomenon's long and complex causal history are explanatorily relevant. That is, the task of the notion of explanatory relevance is to distinguish factors of the type (3) from those of the type (2) – whereas casual relevance distinguishes (2) from (1). In other words, the aforementioned relevance relations *R* constitute different standards or criteria according to which certain causal factors are dismissed as irrelevant and other causal factors are assessed as being relevant to the phenomenon in question and thus are included in the explanation. There have been proposed different criteria for drawing this distinction among causal relationships, for example their stability, their proportionality, their specificity (Woodward 2010), their mutual manipulability (Craver 2007a, 2007b), their activity (Waters 2007), or their substitutability (Sarkar 2005).[56] But the debate is far from being settled yet. What matters to the dispute about explanatory reductionism is whether there are convincing grounds on which certain classes of causal factors (namely, lower-level or higher-level factors) can be dismissed as being explanatorily irrelevant in principle.

Before I continue with this point let me briefly mention a possible objection to the view just presented. Authors like Mackie (1974, 34-36) and Strevens (2008, 41-65; 183f) assume that the notions of causal and explanatory relevance cannot be neatly distinguished. According to their view, causal relations are themselves relative to a certain causal field or to an explanatory framework. Thus, one cannot first identify all causes and then sort out the explanatorily irrelevant causal factors. I do not exclude the possibility that there might be convincing arguments for such a notion of causation and causal relevance. However, this is not the view that is prevalent in discussions about explanatory reductionism in philosophy of biology. In order to assess the arguments that are offered in this debate I therefore take for granted their thesis that causal and explanatory relevance are two distinct issues.

---

[56] Some philosophers even try to identify the one notion of explanatory relevance that holds for all causal explanations (e.g. Woodward 2003). But, as mentioned above, I think it is more promising to pay attention to the different types of causal explanation that are available in scientific practice and to specify different corresponding relevance relations. However, this does not preclude the possibility that there exists a general notion of explanatory relevance that is true for all causal explanations.

Reconsider now the results of Section 2.2.1. There I argued that discussions about explanation that matter to reductionism concern questions about the adequacy of higher- and lower-level explanations. Antireductionists, such as Kitcher, state that higher level factors are explanatorily relevant to higher level phenomena, but lower level factors are not.[57] Reductionists typically deny this and defend the explanatory priority of lower levels or the lowest level of explanation. The refined version of van Fraassen's pragmatic account that I presented in Section 2.2.2 offers a solution to this dispute. This solution goes back to a claim made by Sober. In his paper on "The Multiple Realizability Argument Against Reductionism" (1999) Sober convincingly argues that *generality* (or breadth) and *depth*[58] are two distinct virtues of explanation. He asserts that higher-level explanations have the virtue that they unify what counts as disparate phenomena on a lower level, whereas lower-level explanations are advantageous because they provide more details. He concludes that

> there is no objective reason to prefer the unified over the disunified explanation [i.e. the higher-level over the lower-level explanation]. Science has room for both lumpers and splitters. (1999, 551)

I agree with Potochnik (2009) that higher-level explanations need not be more general than lower-level explanations in the sense that the former have a more limited range of application than the latter.[59] However, this criticism does not affect the general point Sober makes. According to his and my view, neither the reductionist's thesis that lower-level explanations are always adequate, nor the antireductionist's (à la Kitcher or Putnam) claim that they are always inadequate is right. Rather, both higher- *and* lower-level explanations may be adequate. This is the case because both kinds of explanations possess different virtues, namely generality or depth, and these different virtues may be favored in different contexts. In the framework of van Fraassen's refined account of explanation one could say that higher- and lower level explanations are adequate with respect to *different relevance relations*. Higher-level explanations appeal only to causal factors that are located on the same or on a higher level than the phenomenon to be explained, whereas lower-level explanations refer only to causal factors that are located on a lower level than the explanandum phenomenon. Which of these two relevance relations one assumes (higher-level relevance, lower-level relevance, or both) depends on pragmatic factors as well as on how the phenomenon to be explained is specified. Accordingly, there can be no level of explanation that is "right" for all phenomena, irrespectively of the context in which the explanation is given.

---

[57] Some pragmatist-minded antireductionists endorse the more moderate claim that there is not "the right" level of explanation independent from any context.

[58] Note that this is a different notion of depth than the one Woodward and Hitchcock (2003) employ. According to their view, a generalization provides deeper explanations when they are "*more general*" (2003, 198) with respect to hypothetical changes in the system at hand (not with respect to other systems than the system whose behavior is to be explained).

[59] This is why Potochnik argues that they exhibit "different types of generality" (2009, 64).

3 Interim Conclusion

The purpose of this chapter was to reveal the interrelations that exist between the issue of explanation and my account of explanatory reduction. Questions about explanation that are involved in disputes about explanatory reductionism were of interest, too, but only insofar as they affect my analysis of reductive explanation. All in all, the choice of topics addressed in this chapter was quite instrumentalist. I approached only those issues that are relevant to my project of developing an ontic account of explanatory reduction.

The goal of Section 1 was twofold. On the one hand, I introduced different accounts of explanation – namely the CL model (Section 1.1) and the CM model (Section 1.2) – as well as the main critiques that have been raised against them. Thereby I provided the basis, on which in Section 2 I could address more specific questions about explanation that concern the issue of reduction(ism). On the other hand, my aim was to clarify in what sense CM models of explanation are called 'ontic'. This is relevant to my work because I characterize my own account of explanatory reduction in biology as ontic, too (see Chapter V, Section 6).

Let me review the results of Section 1 in more detail. Different accounts of explanation, like the CL and the CM model, provide different answers to the central question of what the nature of (scientific) explanation is. That is, they identify different features that are said to be common to all explanations and to distinguish explanations for instance from pure descriptions. I pointed out that the basic idea of the CL model is that a phenomenon is explained by showing that, given the particular circumstances and the laws in question, the phenomenon was to be expected (Section 1.1.1). In other words, the DN model identifies nomic expectability as the nature of scientific explanation. The objections that have been put forward against the CL model (Section 1.1.2) show that it fails to state sufficient conditions for explanation (because it does not account for the asymmetric features of explanation and fails to exclude irrelevant information) as well as necessary conditions (because not all explanations contain law statements).

In order to avoid objections like these, several authors adopt an alternative view of explanation (namely a CM model), which regards causality as being intimately involved in explanation (Section 1.2). Proponents of a CM model concur that the nature of explanation is causal. They reject the thesis that explanations are arguments and that the explanatory force in explanations arise from rendering the explanandum phenomenon nomically expectable, or from subsuming it under a general law. Rather, what renders a description explanatory is the fact that the causes that lead to or make up the explanandum phenomenon are adequately represented. That is, the basic idea behind the CM model is that to explain a phenomenon amounts to situating it in the causal structure of the world. Different versions of the CM model have been proposed, for instance Salmon's transmission theory (Section 1.2.1) and the mechanistic account of explanation (Section 1.2.2). Also CM models encounter serious objections (Section 1.2.3). In general, it has been alleged that they rely on the concept of causation that has not been sufficiently specified so far, that they provide no account of what it means for a causal factor to be explanatorily relevant, and that they do not account for non-causal explanations in science.

In Section 1.3 I examined what it means to characterize an account of explanation as "ontic". The result of my analysis was that I disagree with philosophers like Craver and Strevens that what makes an account of explanation ontic is the thesis that explanations are objects or facts that exists in the world itself, independently from scientists requesting, discovering, and representing them. Rather, I argued that explanations are epistemic units, and that calling a conception of explanation ontic indicates a certain thesis about what determines whether a description is explanatory or not. Ontic accounts of explanation trace the difference between explanations and non-explanations, ultimately, back to differences in the world, and not to logical features of linguistic entities. That is, the ontic character of a conception of explanation can be linked to the thesis that what determines whether a description is explanatory is whether it truly represents certain causal facts in the world. This result will play a significant role in characterizing my own account of explanatory reduction as ontic, too (see Chapter V, Section 6).

On the basis of these preliminary notes I could then turn to the more specific question of which relations exist between issues of explanation and issues of reduction (Section 2). The central question was whether the debate about explanatory reduction amounts to a debate about explanation (as van Regenmortel claims) or whether questions about explanatory reduction are distinct and can be independently answered from questions about explanation (as Rosenberg states). One of my main theses in this section was that this question must be answered differently with regard to discussions about accounts of reduction (Section 2.1) and with regard to debates about reductionism (Section 2.2).

In Section 2.1.1 I pointed out that the question of what constitutes the reductive character of biological explanations (i.e. the question of explanatory reduction) does not boil down to the question of what characterizes an adequate explanation (i.e. the question of explanation). In other words, in my view, what makes an explanation reductive is different from what makes a description explanatory. I further argued that my account of explanatory reduction remains uncommitted with respect to whether the CL or the CM model adequately captures the nature of biological explanation and that this should be regarded as a virtue rather than as a shortcoming of my account. In order to maintain this neutrality I use the general phrase 'explanatorily relevant factors' to distinguish those entities, properties, interactions, activities, causal processes, regularities, or laws that can be described in the explanans of an explanation (Section 2.1.2).

In Section 2.2 I provided an answer to the question of whether the debate about explanatory reductionism amounts to a dispute about explanation. In order to do so, I first examined which kinds of questions about explanation are at issue (Section 2.2.1). The result was that those disputes about explanation that are relevant to reductionism concern questions about the adequacy of explanations, in particular, questions about the adequacy of higher- and lower-level explanations (i.e. whether certain levels are explanatorily prior to others). Furthermore, I argued that the answers one gives to these questions are affected by one's stance on the pragmatics of explanation. In Section 2.2.2 I showed that Rosenberg is wrong in claiming that disputes about the truth of explanatory reductionism can be kept apart from disputes about explanation. His own defense of Darwinian reductionism demonstrates that disputes about reductionism in fact amount to discussing the adequacy of higher- and lower-level explanations. One of my central theses was that one's stance on this issue is mainly affected by the stance one takes on the pragmatics of explanation (rather than by one's stance on whether CL or CM models are adequate). This is why, in Section 2.2.3, I addressed the question of whether and in what way pragmatic factors influence the adequacy of explanations. I introduced and critically discussed van Fraassen's pragmatic theory of explanation. In order to defend it against Rosenberg's objection that erotetic accounts turn the adequacy of an explanation into a "subjective" matter or into a mere question about the background knowledge and explanatory interests of biologists, I developed a refined version of van Fraassen's pragmatic account. According to my approach, explanation inevitably has pragmatic dimensions. The research context determines which specific phenomenon *P* is to be explained and it affects the question of according to which genuine relevance relation, *R*, certain information is included in the explanans and others is excluded. However, I argued that this does not imply that the adequacy of an explanation is "subjective" or exclusively determined by pragmatics. The reason is that pragmatic factors neither influence which relevance relations at all count as genuine, nor do they influence how these relevance relations are to be specified. On the basis of this refined version of van Fraassen's account I could then reject the thesis that there exists just one right level of explanation for each biological phenomenon (Section 2.2.4). Although the topics I discussed in Section 2.2 concern reductionism and not reduction, I argued that they are nevertheless indirectly relevant to my work. Briefly speaking, the revised version of van Fraassen's pragmatic account of explanation ensures that all different kinds of explanation that can be found in actual biological practice are considered and are not dismissed in advance.

V The Ontic Account of Explanatory Reduction

> *A prevalent… stance in biology is… reductionism, which predicates the study of biological systems at the lowest possible level with the objective of uncovering molecular and biochemical causes.*
>
> (Ana M. Soto/Carlos Sonnenschein 2010, 364)

> *These examples support the importance of being able to think holistically, to look outwardly from the boundaries of the phenomenon under study and thereby place it in a more inclusive context.*
>
> (William Z. Lidicker 1988, 280)

> *[G]ene sequencing and other techniques will soon have isolated all the cell's individual parts and spelled out their isolated functions. Now, it is time to move beyond reductionism. […] Now we need to know how all these things are integrated.*
>
> (Robert F. Service 1999, 81)

Content

The preceding chapters made clear what the main goal of this book is. It is to provide an account of explanatory reduction in biology, that is, an understanding of what it is that makes a biological explanation reductive. How should the line between reductive and non-reductive explanations be drawn and what are the features of biological explanations that determine their reductive or non-reductive character? This project is valuable by itself because the concept of a reductive explanation has not received sufficient philosophical attention so far. Moreover, it is beneficial since it has the potential to advance debates about explanatory reductionism. As I have argued in Chapter II, the philosophical dispute about reductionism in biology is in need of an account that clarifies the notion of a reductive explanation. Any fruitful discussion about the correctness of explanatory reductionism must be based on a clear understanding of what explanatory reduction is. Otherwise misunderstandings and people talking at cross purposes are a daily occurrence. This is not only true for philosophical debates but also, and in particular, for discussions about reductionism in the biological science itself.

Before I start developing my ontic account of explanatory reduction let me sum up where we have come to so far. What crucial insights did my metaphilosophical considerations (Chapter I), my review of the previous reductionism debate (Chapter II), my analysis of recent perspectives on explanatory reduction (Chapter III), and my examination of questions about explanation (Chapter IV) yield that need to be taken into consideration in developing my account? In the following brief summary I will focus on two issues that are of particular importance in this context: first, on the methodological remarks I made in Chapter I on how I proceed in developing my account of reduction. For instance, what are the empirical data on which my account is based? And why is the development of my account better characterized as a critical reconstruction, rather than a mere description? Second, I will pay special attention to those recent claims about explanatory reduction (Chapter III) that constitute notable insights into the character of reductive explanation. These theses will be picked up on in developing my own account in this chapter.

The Methodology of My Account

In Chapter I, I revealed the methodological presuppositions of my analysis. I clarified that my goal is to understand what reductive explanation in contemporary biological practice is (reduction in practice), rather than to propose a view of what reductive explanation ideally should be (reduction in principle). From this it follows that I accept descriptive adequacy as an important criterion of adequacy (which is why my account might be called naturalistic). Moreover, my account is developed in a bottom-up fashion, which means that it emerges from a detailed investigation of the reductive reasoning practices that are present in contemporary biology.

However, in order to be philosophical my analysis must involve taking up a critical stance on the empirical data scientific practice provides us with. In other words, developing a philosophical account of explanatory reduction requires more than a mere description, for instance, of how terms like ‘reductive explanation’ or ‘reductionism’ are actually used in biology. Instead, it calls for a *critical reconstruction* of the empirical data biological practice provides us with. In particular, this means, first, that my analysis does not consider all available cases of reductive explanation, but rather focuses on paradigmatic and important examples of reductive explanations, second, that it explicates assumptions about reductive explanations that are only implicitly present in biological practice, and, third, that in order to construct a coherent and (to a certain degree) universal account I need to assess some of biologists’ statements as incorrect, too vague, or inappropriate, and exclude them from the empirical basis of my account.[1]

The goal of this critical reconstruction is twofold: on the one hand, my analysis aims at identifying *central features* of reductive explanations. That is, it aims at explicating the conditions under which biological explanations succeed or fail to exhibit a reductive character. On the other hand, it is a critical reconstruction also of the *merits and limitations* of reductive explanations, which means that I examine in what way the reductivity of an explanation of certain kinds of phenomena promotes or constrains its adequacy. In short, what is typical for my approach is that it takes biological practice seriously, but nevertheless retains a critical stance towards biological practice.

Finally, let me elaborate on the question of what these empirical data are, on which my analysis is built (see also Chapter I, Section 1.4). In my view, to take biology seriously

---

[1] This critical stance can also be characterized as normative since it involves normative assumptions about which examples should count as paradigmatic or as instances of successful reductive explanations, and which statements of biologists should be regarded as correct or relevant.

implies not merely discussing old standard philosophical examples (like Mendel's second law of independent assortment or the cubical-peg-square-hole example) again and again. Instead, we need to analyze the various reductive reasoning practices that are present in actual biological research and the way they are conceptualized (if they are) in biological discussions about reductionism. More precisely, the empirical data that are involved in reconstructing central features, advantages, and constraints of reductive explanations can be divided into two groups: first, I analyze the characteristics of paradigmatic and important *examples of reductive explanations* from various biological disciplines.[2] These examples are in part those that are discussed in philosophy of biology. However, the majority of examples are taken from the biological literature itself.[3] Second, I develop my account of explanatory reduction on the basis of explicit *discussions about reductionism* that can be found in current biological research papers. Although these biological discussions about reductionism have only received little attention in the philosophical debate about reduction in biology so far, I think they are of great value. In order to get a more comprehensive picture of the kind of discussions I appeal to, consider these three general characteristics of the biological literature on reductionism.

First, if biologists reflect on their work and on the work that is done in their discipline, the issue of reduction is frequently discussed. However, most biological discussions are concerned with *reductionism*, not with reduction. They argue about whether certain fields of research can be accused of being too reductionist or whether in a certain discipline a general trend away from reductionism and towards more holistic approaches can be observed. But this focus on questions about reductionism does not preclude that one can derive relevant insights about reduction from these disputes.

Second, biological discussions about reductionism focus on *methodological* issues or, more precisely, on the fruitfulness or permissibility of applying reductive methods in order to investigate the phenomenon or class of phenomena in question. Frequently these discussions about reductive methods also involve issues concerning reductive *explanations*. However, discussions about reductive explanations are not concerned with the concept of a reductive explanation itself (which is the focus of my analysis). Instead, biologists dispute

---

[2] Since I want my account to apply to biology in general and not to be restricted to certain biological disciplines, the strategy Sarkar (1998), Hüttemann and Love (2011) chose, namely to focus on the analysis of examples from only a few biological fields (e.g. genetics, molecular biology or developmental biology), is not a viable option for me.

[3] For example, they are taken from biological research papers that explicitly concern the issue of reduction(ism) or related topics, such as part-whole explanation, decomposition, etc. I admit that the choice of paradigmatic and important examples of reductive explanations in biology sometimes might be affected by a certain *pre-concept of reduction* one (explicitly or implicitly) endorses – for example, that reduction leads to simplification, that reduction is closely related to mechanisms and part-whole explanation, that reduction involves showing that something is nothing over and above something else, etc. It seems to me that this is not bad – as long as the influence of this pre-concept is not too strong, as long as this influence is revisable in the light of important insights about what scientists treat as reductive explanation, that is, as long as this does not result in imposing an ill-fitting view of reduction on biological practice.

whether reductive explanations for certain phenomena are adequate or not. That is, they presuppose a certain understanding of what makes an explanation reductive but, even if they make this understanding explicit (which is rarely the case), they do not argue about it. Thus, there is philosophical work left to do, namely to figure out what biologists mean by a reductive explanation when they quarrel about the adequacy of reductive explanations or about the reductionist approach in general.

Third, the overall tone of the current discussions about methodological and explanatory reductionism in the biological sciences is *negative*. The argument is that, especially when it comes to the behavior of complex systems and to complex diseases, reductive research strategies are either of limited value (e.g. Gallagher and Appenzeller 1999; Ahn, et al. 2006a) or completely misleading or "ineffective" (Soto, et al. 2009, 5) because they provide deficient results and, if applied exclusively, they result in inadequate reductive explanations. Hence, many contemporary life scientists challenge the normative thesis of methodological reductionism that research should always be pursued in a reductionist manner (see Chapter II, Section 3.2). Likewise, they question the truth of explanatory reductionism by pointing out that many phenomena cannot be adequately explained in a reductive manner. However, it is important to note that, despite their overall negative attitude, most scientists concede that reductionist research strategies have been and still are of value and that the behavior of many systems can adequately be explained reductively. One can read statements like "r[]eductionism has served […] biologists well in the past decades" (Powell 2004, 299), or "[o]ver the past century, significant advances in medical practice and healthcare have been achieved based on traditional reductionist biomedical research" (Roukos 2011, 695). In general, criticisms of reductionism in the life sciences tend to point out only that the value of reductive methods and the adequacy of reductive explanations is *limited* (Fang and Casadevall 2011, 1402). For instance, reductive explanations are only adequate with regard to a limited class of phenomena like acute, simple diseases (e.g. urinary tract infection, appendicitis, etc.; Ahn, et al. 2006b, 2). Or reductive strategies provide only limited insights into a complex system, that is, insights into the behavior of the isolated parts but not into the "dynamics of a system as a whole" (Sorger 2005, 9). What exactly the merits and limitations of reductive explanations are will be pointed out in the course of this chapter.

Gained Insights on Reductive Explanation

The central task of Chapter II was to introduce the longstanding debate about reduction(ism) in the philosophy of biology. I identified four major lessons that one should learn from the previous debate. These lessons also served to state more precisely what the central target of my analysis is. I do not aim at defending (anti-)reductionism, but at developing an account of reduction in biology. In doing so, I am not concerned with ontological reduction and not primarily with methodological reduction. In particular, I do not try to advance a Nagelian model of theory reduction. Rather, I break fresh ground and develop an understanding of what *explanatory reduction* in biology is.

In Chapter III I discussed the philosophical work that has been carried out on the topic of explanatory reduction so far. I distinguished two perspectives: first, Rosenberg's view of reduction as a relation between two explanations, namely between a higher-level and a lower-level explanation of the same phenomenon; second, Sarkar's, Hüttemann's, and Love's work on the reductive character of individual explanations. Various problems with Rosenberg's understanding of explanatory reduction (see Chapter III, Section 1.3) led me to discard the first perspective and to join Sarkar, Hüttemann, and Love in their search for the determinants of the reductivity of biological explanations. My examination of Sarkar's, Hüttemann's, and Love's accounts yielded crucial findings with regard to what are (and what are not) major features of reductive explanations in biological science. Since these findings make a significant contribution to developing my own account of explanatory reduction, I briefly repeat the most important of them here.

First, I argued that, in line with Sarkar's, Hüttemann's, and Love's perspective, explanatory reduction must be interpreted as a relation that holds between the two parts of an explanation, namely between the description of the explanandum phenomenon and the description of the explanatorily relevant factors that are referred to in the explanans.

Second, it has been argued that reductive explanations are identical to other kinds of explanation, namely to mechanistic explanations (Wimsatt 1976a), part-whole-explanations (Kauffmann 1970; Hüttemann/Love 2011), and lower-level explanations (Sarkar 1998; Hüttemann/Love 2011). Although I agree that these kinds of explanation are closely related I will show that it is misleading to identify reductive explanation with part-whole or mechanistic explanation. Even if the set of all reductive explanations overlap to a great extend with the set of all part-whole explanations and of all mechanistic explanations, I think it is important to note that they do not coincide (see Section 5).

Third, Sarkar (1998) argues that an appropriate analysis of reductive explanation should focus on substantive issues, not on formal ones (see Chapter III, Section 2.2.1). This is a crucial insight that stands behind much of the criticism that can be raised against Nagelian models of theory reduction (see Chapter II, Section 4). In accordance with Sarkar's focus on substantive issues, I develop an account of explanatory reduction that can be characterized as "ontic" because it traces the reductive character of an explanation back to the fact that it appeals to certain relations and facts that exist in the world (see Section 6).

Fourth, Sarkar, Hüttemann, and Love propose different criteria or aspects of reductive explanations in the biological sciences. That is, they provide an answer to the central question that I am concerned with, too, namely what are the features that reductive explanations have in common. Sarkar (1998) identifies three criteria of reductivity, namely fundamentalism, abstract hierarchy, and spatial hierarchy. Hüttemann and Love (2011) highlight three aspects of reductive explanations, namely intrinsicality, fundamentality, and temporality. In Chapter III (Section 2.2.4 and 2.3.4) I pointed out several shortcomings Sarkar's, Hüttemann's, and Love's analyses display. However, they comprise notable insights, too, from which my own analysis can profit. For instance, the "spatial hierarchy" and "fundamentality" criteria or aspects express the finding that in reductive explanations the factors that are referred to in the explanans are located on a lower level than the phenomenon to be explained (see Section 2). Furthermore, Hüttemann and Love are the first authors who recognize a difference between the claim that something is located on a lower level (fundamentalism) and that something is located inside the spatial boundary of a system (intrinsicality). This difference will be taken up and further elaborated in my analysis (see Section 3).

Chapter IV addressed the question of what the *view of explanation* is that underlies my account of explanatory reduction. I argued that it is important to keep apart the question of explanation from the question of explanatory reduction. The former calls for criteria that allow distinguishing adequate explanations from non-explanations, whereas the latter calls for standards that help to draw the line between two classes of adequate explanations, namely reductive and non-reductive adequate explanations. I pointed out that my account of explanatory reduction answers only the latter, but not the former question. That is, it remains neutral with respect to the old dispute whether nomic expectability (CL model) or causation (CM model) is most important to scientific explanation. In order to preserve this neutrality, I speak of explanatorily relevant "factors" that are referred to in the explanans, not of laws or regularities or causal factors. However, my account remains not completely neutral with respect to questions about explanation. Certain views about explanation (such as Rosenberg's non-erotetic view) need to be ruled out because they would narrow the set of explanations which count as reductive in an inappropriate way. To prevent this, I defend a refined version of van Fraassen's pragmatic account of explanation.

This is where we have gotten so far. On this basis I can now develop my own account of explanatory reduction in biology. I will proceed in two major steps. The first part of this chapter (Section 1) serves to specify four concepts that occupy center stage in my analysis of the features of reductive explanations, namely the concept of a biological *system*, the concept of its *context* or environment, the concept of *parts* of a system, and the concept of *levels* of organization. The specification of these concepts is essential because they are central to answering my main question of which features of biological explanations determine their reductive character. This becomes apparent if one considers the answer I give to this question. My central thesis in the second part of this chapter will be that reductive explanations in biology exhibit three major characteristics: first, they refer exclusively to factors that are located on a lower level than the system whose behavior is to be explained, second, they focus on factors that are internal to (i.e. parts of) the system in question and ignore or simplify environmental factors, and third, they appeal only to parts in isolation. This answer is convincing only if it is not left completely unclear what a biological system is, how a system is distinguished from its environment, under which conditions an entity counts as a part of a system, and what a level of organization is. Accordingly, the goal of the first part of this chapter is to specify the concept of a system and its context (Section 1.1), to clarify what it means to decompose a system into its parts (Section 1.2), and to reveal which notion of level my account of explanatory reduction presupposes (Section 1.3).

On the basis of these conceptual clarifications I can then in the second part of this chapter (Sections 2 to 6) develop my account of explanatory reduction in biology. I will do so by analyzing, on the one hand, paradigmatic and important examples of reductive explanations from different biological fields and, on the other hand, discussions about reductionism in the contemporary biological literature. The result of this analysis will be that reductive explanations in the biological science possess three characteristics (two of which are necessary conditions, one of which is a typical feature that most reductive explanations exhibit, and all of which taken together might be a sufficient condition for reductivity). I will argue that reductive explanations display a lower-level character (Section 2), focus on factors that are internal to the system of interest (Section 3), and describe only the parts of a system in isolation (Section 4). I conclude by pointing out the similarities and differences between reductive explanations and part-whole explanations and between reductive explanations and mechanistic explanations (Section 5). Furthermore, I explain why I call my account an *ontic* account of explanatory reduction (Section 6).

1 Preliminary Remarks: Systems, Contexts, Parts, and Levels

The findings of the previous chapters (especially those presented in Chapter II, Section 3.3 and in Chapter III) suggest the assumption that explanatory reduction has something to do with three tasks: first, with identifying biological systems by separating them from their context (or environment), second, with decomposing a system of interest into certain parts, and third, with identifying levels of organization and assigning entities of interest to particular levels. This is just a preliminary assumption, which will be specified and justified in the course of this chapter. However, it supplies us with enough reasons to carefully examine the concepts that are involved in this assumption. Accordingly, the goal of this section is to clarify the concept of a biological *system*, the concept of the system's *context* or environment, the concept of a *part*, and the concept of a *level of organization*. In Sections 2, 3, and 4 I then show in which sense these concepts are of particular importance for specifying the central features of reductive explanation in biology.

1.1 The Concept of a Biological System

Each biological investigation begins with identifying a particular object of study and, in most cases, with posing one or more specific questions about this object. Frequently, the object of study is a biological system, for instance, a particular kind of cell or cell complex, an organ or organ system, a particular organism, or even a whole population or ecosystem. Biological systems can be found on (almost) all levels of biological organization (see Section 1.3). At the outset of any research the biological system that is to be investigated needs to be *individuated*. The task of individuating the system of interest involves for example separating it from everything else that does not belong to the system. This "everything else" is called the *context* or *environment*[4] of a particular biological system. To individuate a system by deciding which entities and interactions are parts of this system and which belong to a system's context is common practice in biology. However, the philosophical questions it raises are less easy to answer. How can the notion of a biological system be specified? That is, in what circumstances does a bulk of entities add up to a biological system? And according to which criteria do biologists decide whether something is a system, respectively whether an entity belongs to a system or to the environment of that system? In other words, how is the boundary between a particular system and its context to be drawn? These questions touch important metaphysical and methodological issues, which cannot be exhausted in this book. However, in what follows I present some ideas concerning the individuation of biological systems that, at least, can serve as stimulating starting points to approach the questions I just posed.

Let us start with the question of how the concept of a biological system can be specified. Much philosophical (and biological) work has been carried out to explicate the notion of a biological organism (e.g. Wilson 2005, 2008; Dawkins 1982; Folse/Roughgarden 2010; Dupré 2010; Pradeu 2010). But only few philosophers have given attention to the broader category of a biological system.[5] One reason for this could be that the concept of a biological system is notoriously vague and frequently used as a filler term for more precise concepts, such as organism. However, this does not imply that nothing insightful can be said about the notion of a biological system. To the contrary.

In my view, what is crucial for biological systems is that they consist of a set of entities (that exhibit particular properties, are organized, and interact with each other in a specific way) that together form a *unified whole*. More precisely, this means that biological systems exhibit at least two (necessary) characteristics.[6]

---

[4] This is a very broad notion of environment, which means something like Brandon's "external environment" (1990, 47-49) but is not restricted to the environment of organisms (since the category of biological systems comprises, but is not confined to the class of organisms).

[5] Of course, remarkable work in systems theory has been done on the general notion of a system (including social and political systems, climate systems, non-living systems, etc.). And more recently philosophers have engaged themselves in discussions about systems biology.

[6] It may be argued that biological systems are individuated according to a *third condition* (that is closely related to the first): biological systems are unified wholes because they consist of such entities that exhibit the strongest and most frequent interactions to other parts of the system, rather than to entities that belong to the system's environment. For a related argument see Simon 1962, 477; Wimsatt 1974, 72; 2007, 184, 204, 209.

> Biological systems consist of a set of interacting entities (i.e. its parts) that
> (i) *work together* to bring about those functions[7]/behaviors that the system characteristically displays, and that
> (ii) are located within a defined zone, which is surrounded by a *spatial boundary*.

In other words, only those entities and interactions are parts of a system (and not of its environment) that, first, contribute to bringing about one or more behaviors that are regarded as being typical for this system (i.e. "what the system is doing"; Kauffmann 1970, 259) and, second, that are contained within the spatial boundary of the system. This way of understanding the notion of a biological system brings up two new questions: first, what are these characteristic functions or behaviors that determine whether an entity belongs to a system or not? Second, what is the nature of the spatial boundary that demarcates the system from its context? I approach these two questions in the following sections.

Before I do so, let me add a crucial remark. I just claimed that biological systems are individuated by performing two tasks, namely by identifying those functions or behaviors that are characteristic for the system (e.g. the distribution of blood and the transport of nutrients as important functions of the circulatory system of humans) and by drawing a spatial boundary between system and environment (e.g. the membrane that surrounds the circulatory system). However, my above comments also reveal that the individuation of a biological system is intertwined with the decomposition of the system into parts (see Section 1.2). This is the case because to identify a system involves deciding which entities belong to the system, that is, are *parts* of the system, and which are not. This does not imply that one cannot individuate a system before having completely decomposed it. In fact, this would be impossible because decomposition presupposes having individuated the system (in particular, it presupposes having identified a behavior of the system according to which the system is to be decomposed). Instead, it means that the individuation of a system already involves decisions about what are parts of a system and what are not, and that the two tasks are not completely independent from another, but rather intertwined. The following figure illustrates this:

---

[7] In this context, the notion of a function of a system can best be understood as a certain *causal role* the system plays in its context (e.g. Cummins 1975). For instance, the immune system, which constitutes a subsystem of an organism, plays a prominent role in protecting the organism against disease. However, with this I do not want to claim that etiological accounts of functions, which link the concept of a function closely to the principle of natural selection, are completely misleading. Rather, I feel attracted by pluralistic approaches that assign an important role to both accounts of functions (like the one Kitcher 1993 develops).

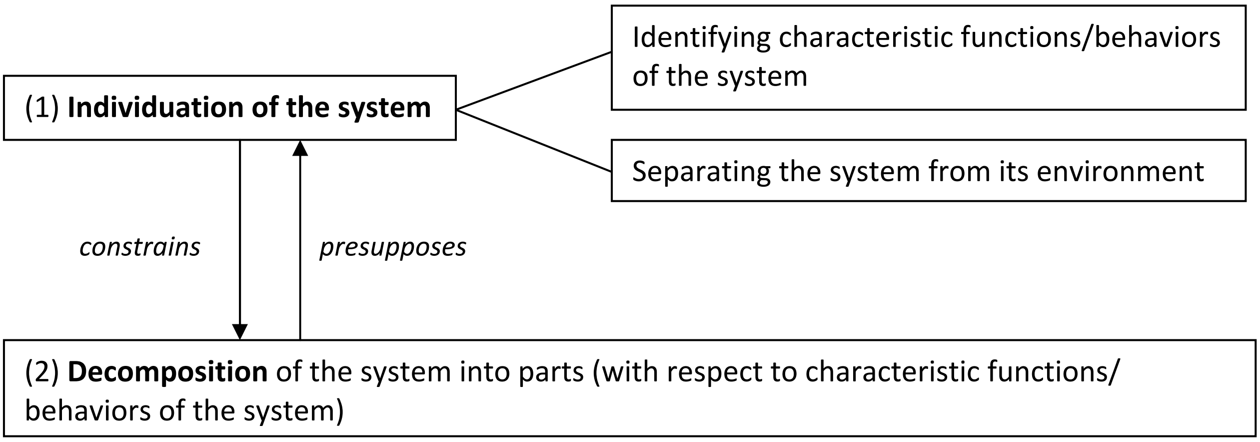

Figure V.1: Individuation and Decomposition of Biological Systems

1.1.1 Identifying the Behavior(s) of a System

In the foregoing section I argued that biological systems are constituted by a set of organized and interacting entities that meet at least two conditions: (i) they work together to bring about those functions/behaviors that the system characteristically displays, and (ii) they are located within a defined zone, which is surrounded by a spatial boundary. With regard to the first criterion an opponent of my proposal could raise the following objection: biologists typically decompose systems by selecting a certain behavior of that system and determining which entities and interactions are relevant to bringing about this behavior (e.g. Kauffmann 1970, 259f; Glennan 2002, 344; Craver 2007a, 122-128; see also Section 1.2). Accordingly, the first criterion (i) should not refer to different functions or behaviors of the system, but rather should distinguish *one singular* behavior according to which the system and its parts are to be individuated. That is, an opponent could propose the following revised version of the first criterion, let us call it (i*):

> Biological systems consist of a set of interacting entities (i.e. its parts) that
> (i*) work together to perform one particular characteristic function/behavior

For the sake of the argument let us accept this revision and see where it takes us. Consider an example of a biological system, namely a cell. One thing that is important to a cell is its ability to divide into two or more daughter cells, a process called cell division. However, not all entities that are contained within the spatial boundary of the cell, namely the cell membrane, are relevant to bringing about the cell division. That is, there exist entities (like peroxisomes or the Golgi apparatus) that meet criterion (ii), but fail to satisfy criterion (i*). This gives rise to the implausible thesis that entities such as peroxisomes or the Golgi apparatus do not belong to the system of a cell.

In order to avoid this implausible consequence my proposal is to regard not only one, but *all* behaviors that a biological system characteristically displays or is disposed to display as crucial to the individuation of this system.[8] Of course, it is still possible that biological systems exhibit only one typical behavior or function, according to which they are individuated. Whether and for how many systems this is true depends on how fine-grained an individuation of systems you seek and on how fine-grained a description of behaviors or functions you make. For instance, a cell, in general, displays much more disparate behaviors than a Helper T cell. Additionally, the function or behavior of the circulatory system of humans can be described, on the one hand, quite broadly as passing different entities to and from cells in the body and, on the other hand, in a more fine-grained way that differentiates between the function of distributing nutrients (such as amino acids, electrolytes, and lymph) and the function of transporting gases and blood cells.

The decision to treat *all* behaviors (and not only one) that a biological system characteristically displays as being essential for the individuation of this system introduces an important *difference* between *systems and mechanisms*. According to the New Mechanists, a mechanism is always a mechanism *for* one particular behavior or phenomenon (e.g. Glennan 2002, 344; Craver 2007a, 123; Bechtel 2006, 29).[9] One reason for this is that the concept of a mechanism is closely linked to the concept of a mechanistic explanation. A mechanism is viewed as “a set of entities and activities organized such that they exhibit the phenomenon to be explained” (Craver 2007a, 5). That is, mechanisms tie only those entities and interactions (or activities) together, which contribute to bringing about the *specific* behavior of a system that is *to be explained*. From this it follows that in the majority of cases mechanisms constitute only *subsets* of biological systems. To put it another way, not all entities and interactions that are part of a system are also part of a mechanism for one of the system’s behaviors (more on this in Section 1.2.1 and 1.3.3). For example, the mechanism of protein synthesis consists of the genome of a cell, its nucleus, r- and t-RNA, ribosomes, free amino acids, and a whole bunch of other proteins, all performing certain activities and all organized in a specific way (see Alberts, et al. 2008, Chapter 6). Although the list of interacting entities that are part of the mechanism for protein synthesis is already quite long, the biological system cell comprises far more entities than them (e.g. centrioles, lysosomes, peroxisomes, Golgi apparatus, mitochondria, etc.). Hence, even if the mechanism for protein synthesis may have the same outer boundary of the corresponding system, namely the cell (which might be questioned; see Section 1.3.3), it contains only a special subset of those entities and interactions that are parts of the cell. That is, the mechanism for protein synthesis is only a subset of the cell – spatially and temporally regarded. A helpful illustration is to think of a mechanism as a *Swiss cheese*, whereas a system will be a cheese without holes (or an assemblage of islands; see Section 1.1.2).

---

[8] Bechtel and Richardson hold a similar view: “Before it is possible… to develop a fully mechanistic explanation… it is necessary to identify *what functions* are performed and what system performs these functions. We speak of this as isolating the locus of control.” (2010, 35; my emphasis)

[9] Commonly, these behaviors or phenomena are types (with varying range), not tokens.

1.1.2 Drawing the Line between System and Context

The parts of a biological system are not only bound together because they are all relevant to the system exhibiting certain behaviors or functions. I have argued above that, in addition, biological systems form a unified whole because they are all located within a defined zone that is surrounded by a certain *spatial boundary*. For instance, cells are delimited from their context by the cell membrane. Similarly, the human circulatory system is enclosed by a membrane that separates it from its context, including other adjacent systems, such as the respiratory or the excretory system. Organisms, too, are said to be surrounded by an outer surface (and, depending on one's point of view, also by an inner surface like the inside of the gastro-intestinal track). Examples like these support the central idea that biological systems are delimited from their environment by a spatial boundary, which ideally is a *continuous, constant surface*. Each entity that is located *inside* this boundary (including any interaction that takes place between these entities) is referred to as 'part of the system' and as 'internal'. Everything located *outside* this boundary is called 'part of the environment' and 'external'.

This is a catchy idea. But unfortunately, the biological world is not always that simple. Some biological systems cannot be so neatly separated from their environments. In what follows I focus on three kinds of delineation-problems.

First, there exist biological systems whose boundary is far from being one *continuous* and *unvarying* surface. Consider, for example, a Golgi apparatus. This is a cell organelle that is part of the endomembrane system of the cell. It is composed of several stacks of membrane-enclosed structures, known as cisternae. The Golgi apparatus serves to modify, sort, and package macromolecules for cell secretion or use within the cell. In order to do so, transport vesicles continuously bud off from one compartment of the Golgi apparatus and fuse with another, or they are transported to other cell organelles and fuse with their membrane.

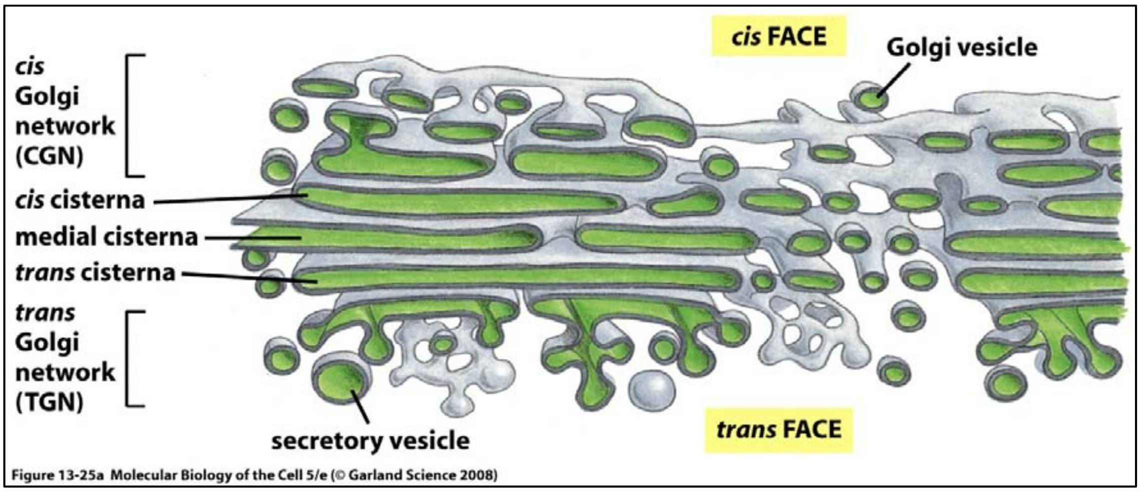

Figure V.2: The Golgi apparatus (Alberts, et al. 2008, 771)

The challenge that this example poses to the idea that biological systems are surrounded by a continuous, unchanging surface is twofold: on the one hand, the boundary of the Golgi apparatus is not unchanging, but in a *constant flow*. Continuously, transport vesicles bud off from one compartment of the Golgi apparatus and fuse with another. To make it even worse, at one time a part of the membrane and certain molecules that are located next to this part of the membrane belong to the Golgi apparatus. At another time this part of the membrane buds off, the vesicle containing certain molecules is transported away, it for instance fuses with the cell membrane, and releases the molecules. The part of the membrane and the molecules that previously were parts of the Golgi apparatus, and respectively constituted its outer boundary, are not parts of the Golgi apparatus anymore. On the other hand, the boundary of the Golgi apparatus is far from being one continuous surface. Even if we took a snap shot of the Golgi apparatus it would be surrounded by several *disconnected* boundaries, rather than by one continuous surface. The reason is that the Golgi apparatus is composed of several stacks of cisternae and vesicles, that are (at most times) not spatially connected with each other.

The latter problem arises when one tries to individuate a population, too. Consider, for instance, a population of black-headed gulls (*Chroicocephalus ridibundus*) breeding on the coast of West Europe. The individual members of a population are not physically connected, but rather distributed over a certain spatial area. Where does the spatial boundary of the system run? If one draws one line round all individual black-headed gulls that belong to the population other organisms will be included that clearly do not belong to the population – for instance, birds of other species that fly between the gulls, or small crustacean or scallops that the gulls search for to eat. Consequently, this is not a viable option. Instead, I suggest conceiving the boundary of a population as the *totality of all disconnected boundaries* of the individual members of the population. Likewise, the boundary that demarcates the Golgi apparatus from its context (to a particular time *t*) is the sum of all disconnected boundaries of the cisternae and vesicle that compose the Golgi apparatus (to the time *t*).[10] Abandoning the ideal that a biological system *must* be surrounded by one continuous spatial boundary appears to be a necessary step in order to account for biological reality. An entity could still be referred to as 'internal' in case it is located inside of one of these several enclosed spaces and as 'external' in case it is located outside of these different boundaries.

Second, other examples raise the problem of how to distinguish a biological system from other biological systems. This is particularly true with respect to a specific kind of biological system, namely with respect to organisms. There are several examples that show how complicated it actually can be to draw the line between an organism and its

---

[10] As this statement shows, the problem that the boundaries of some biological systems vary over time (as in the case of the Golgi apparatus) can be solved by adding that the identification of a certain boundary is relative to a certain point or period of time.

environment, which includes other organisms. For instance, in case of symbiotic associations in is unclear whether we are dealing with one or with more organisms. Consider for instance a lichen, which is a symbiotic association between a fungus and a photosynthetic microorganism (usually either a green alga or cyanobacterium).

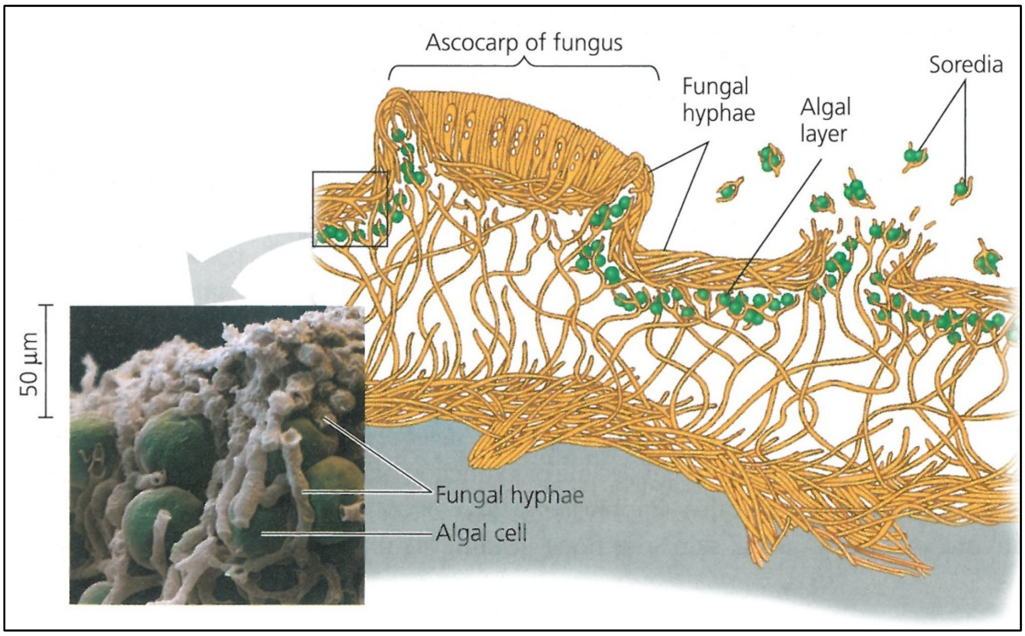

Figure V.3: Anatomy of an Assomycete Lichen (Reece, et al. 2011, 931)

On the one hand the biologists speak about the lichen as a composite, but discrete organism. On the other hand the characterization of a lichen as a symbiotic association suggests that the object of study consists of more than one separate organism, namely of a fungus and an algae. A similar example is a living coral reef, which consists of coral polyps and zooanthellae that have formed a deep, symbiotic relationship over evolutionary time. Superorganisms like social insect colonies (e.g. ants, wasps, bees, and termites) are extremely challenging cases, too. On the one hand, it seems to be clear that insect colonies are groups of organisms (more precisely, “trait groups”; Wilson 1975). On the other hand, it can be argued that insect colonies are to be viewed as a single organism, namely as a “superorganism” (Wheeler 1920), because natural selection seems to be acting on the insect colony as a whole. Examples like symbionts and superorganisms reveal that it may be difficult to decide whether other organisms should be included within the boundary of an organism or not. Contrary to this, cases of “extended phenotypes” (Dawkins 1982) even suggest that previously “non-organismic” parts of the environment should be viewed as belonging to the organism in question. For instance, one might argue that the house construction is a part of the Caddisfly larva because it is necessary for the survival of larva and spatially closely bound to the larva.[11]

---

[11] For further reading see, for instance, Sterelny/Griffiths 1999, Chapter 3 and 8 and, of course, Dawkins 1982.

Third, one could raise an objection that does not concern the delineation of the boundary between system and environment itself, but the subsequent classification of certain factors as being internal and others as being external to the system in question. Since classifying factors as internal or as external presupposes locating them inside or outside of the spatial boundary or boundaries of the system, it is only applicable to factors that have a definite *spatial localization*. This makes it difficult to decide whether such factors as temperature or pressure are internal or external to a certain biological system. One could try to overcome this problem by talking for example about the temperature of certain membrane-enclosed spaces like a cell, which can be distinguished from the temperature of the environment that surrounds the cell. Alternatively, one can just accept that there exist a *few* kinds of factors whose classification as internal or external turns out to be problematic.

In conclusion, the considered examples show that it might sometimes be challenging to separate a biological system from its environment. If this undertaking fails altogether it may be that the biological system cannot be sufficiently identified. As a consequence, it might be impossible to perform the subsequent tasks of decomposing the system into parts and identifying different levels of organization (see also Section 1.2.4). However, this is the worst case scenario and there is every indication that cases like these are very rare. In the majority of cases the idea that a biological system can be separated from its environment by identifying one continuous boundary (that may change through time) is appropriate. And even if there is not the one continuous surface, the boundary of a biological system can still be regarded as a conglomerate of several disconnected boundaries, and thus be clearly distinguished from its environment. Also examples like symbiotic associations, superorganisms, and extended phenotypes do not show that biologists *fail* to perform the task of distinguishing system and environment. They just show that biologists do not always agree on how to draw this line. Moreover, examples like these suggest that sometimes there might be more than one way of partitioning the world into systems, and so of separating a system from its environment (see also Wimsatt 2006a, 2007).

1.1.3 Multiple Boundaries?

Wimsatt recently has claimed that "[a]nalyzing complex systems often requires simultaneous use of different decompositions, boundaries, and contexts" and that this is "not just confused thinking, but a breakdown of concepts of well defined objects". He continues to argue that in biological investigations the system boundaries must be placed "out far enough to include all relevant parts of key mechanisms" and that "the boundaries required to explain different aspects of the behavior of a part of… a system may differ" (2006a, 461). Although Wimsatt has a point when he emphasizes the variability of decompositions and of drawing system boundaries, I think his statement is, at least with respect to system boundaries, too radical. It portrays the individuation of biological systems (which includes identifying their boundaries) as a highly flexible and even arbitrary matter.

This picture neglects that the boundaries of biological systems often correspond to structures (such as membranes) that exist in nature, which is why biologists are quite constrained in how they draw the boundary of a system. Let me explain this in detail.

To begin with, Wimsatt's advice that the boundaries of the investigated system should be placed out far enough to include all relevant factors is misleading. Assume we have a case in which changes in the environmental conditions are so important to the behavior of a system that they cannot be ignored or treated as constant. In cases like these you cannot simply expand the boundary of that system so that relevant environmental conditions are included in the system.[12] To be exact, you *can* redraw the boundary. But this amounts to drawing the boundary of a *different system* than the original one.

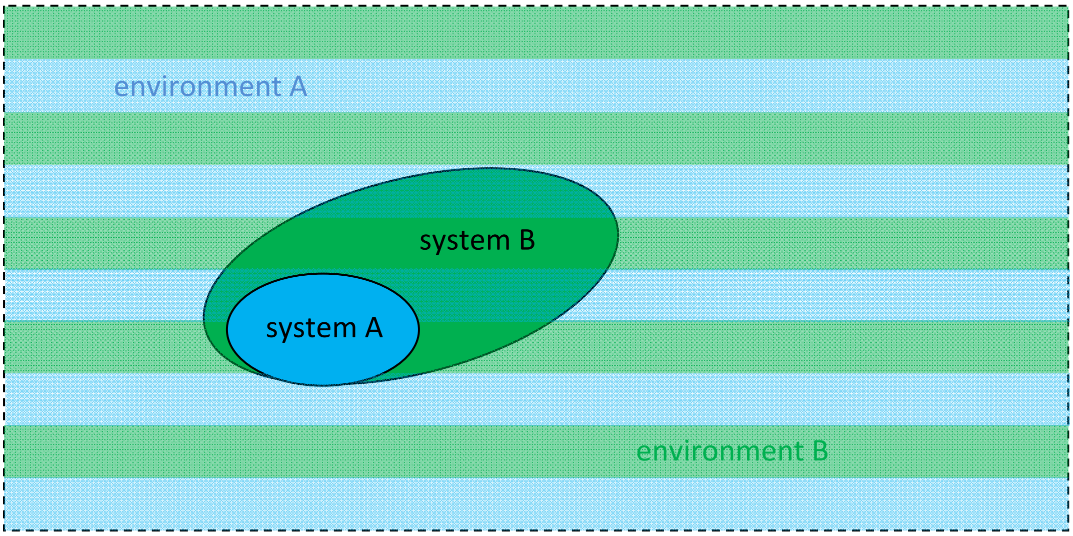

Figure V.4: Redrawing the Boundary of Systems

Wimsatt confuses this since he talks about expanding a system's boundary so that it encompasses "aspects of its environment" (2006a, 461). This suggests that one can simply expand the boundary of a system and turn previously environmental factors into parts of the system. However, it seems to me that in most cases this is not possible since the redrawing of the boundary of a system in such a way that it includes parts of the environment amounts to drawing the boundary of a different system. The intuition that underlies this claim is that it matters for the identity of a particular system where exactly its boundary runs (i.e. what its parts are). This does not preclude that biologists may frequently modify where they place boundaries (which means that they change their object of study). But it precludes that biologists vary greatly about where they draw the boundary of one *particular* system. In other words, Wimsatt may be right that different explanatory interests require drawing different boundaries. But in most cases this means that biologists study different systems, not that they study one system and constantly redraw the boundary of that system.

---

[12] But this is exactly what Wimsatt suggests: "reductionists frequently must expand the boundaries of what was originally taken as the system for analysis to include structured aspects of its environment" (2006a, 461).

Separating a biological system from its environment by identifying its boundary is far from being a highly flexible or arbitrary matter (although it sometimes might be a *difficult* enterprise; see Section 1.1.2). Where the boundary is to be drawn depends, on the one hand, on the question of which parts are relevant to the behaviors or functions that a system is said to display (i.e. on decisions about what the parts of a system are; see Section 1.2). On the other hand, the boundaries of most biological systems fall together with certain structures that exist in nature, for instance, with membranes that dissect organisms into certain segments or that demarcate whole organisms from their environment. Frequently, the boundaries of biological systems involve also qualitative heterogeneity (of material constitution, texture). For instance, the inside of *Euglena*, a protist, contains many cell organelles, macromolecules, and ions that cannot be found in its environment (pond water). Another example is the mammalian cardiovascular system, which contains blood that cannot normally be found in other parts of the organism. In sum, this supports the assumption that the boundaries of biological systems are to be characterized as "*bona fide boundaries*", that is, as boundaries that exist in nature, rather than as "*fiat boundaries*" (Smith/Varzi 2000, 402), which means boundaries that are introduced through human demarcation.[13]

1.2 Decomposition of Systems into Parts

In Chapter II, Section 3.2.2 I introduced decomposition as the reductive method *par excellence*. In the biological literature it is also referred to as "dissection of biological systems into their constituent parts" (van Regenmortel 2004b, 1016), as "analysis" (Mayr 1988, 475), or as "downward looking"[14] (Lidicker 1988, 278; Byerly 2003, 337). Decomposition is not the only reductive strategy that is used in biological science (see Chapter II, Section 3.2.2), but it is the most prominent one. What is more, the method of decomposition is of particular importance for the reductive character of biological explanations. (Why this is so will become apparent in the course of this chapter).

Decomposition is the strategy by which a biological system is divided into its *parts*. For instance, macromolecules are dissected into subunits that are arranged in relation to each other in a specific way, populations are carved into individual organisms, and ecosystems are divided into species that interact with each other in a particular manner. The converse strategy of putting the parts together again (which is often a much more challenging task) is also known as "recomposing" (Bechtel/Richardson 2010, xxxvii) a system.

The relation between parts and wholes is called *composition*. Although it may be connected with the relation of *constitution* (see Chapter II, Section 2.1.1) the two notions must be kept clearly apart. Composition is the relation between something and its parts, whereas constitution is the relation between something and what it is made of (Evnine 2011,

---

[13] Of course, this is not to say that the world completely determines where to draw the boundary of a system.

[14] Characterizing decomposition as "downward looking" presupposes a notion of levels that traces the difference between levels back to compositional relations between parts and wholes (see Section 1.3).

212f). For instance, an animal cell is composed of a nucleus, of mitochondria, of an endomembrane system, of the cytoskeleton, of the cell membrane, and of several other parts. Each of these entities stands in the relation of composition to the cell since each of it is a part of the cell. Contrary to this, the relation of constitution holds only between the entire set of *all* parts of the cell, which are spatially and temporally *organized* in a certain way, and the cell. The following figure illustrates this difference:

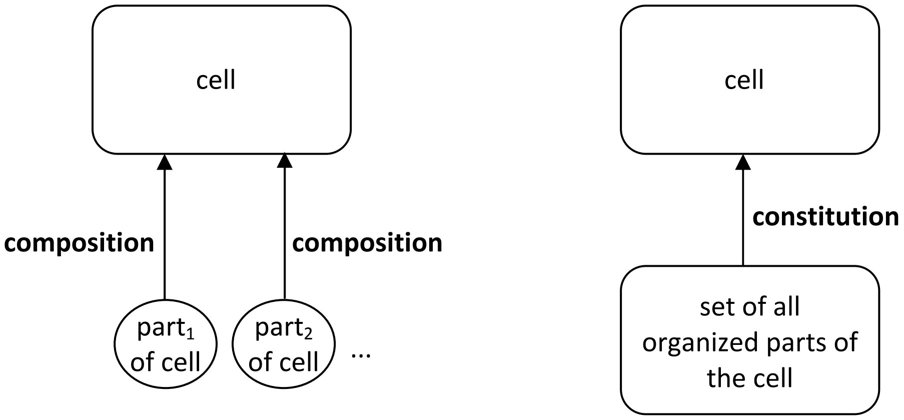

Figure V.5: The Difference between Composition and Constitution

This, however, does not provide an exhaustive answer to the central question of what it means to "be a part of" a biological system, respectively to "compose" a system. In order to answer this question in the next section I distinguish three different views about parthood, and side with one of them, namely with the view that parts of biological systems are genuine parts (Section 1.2.1). However, this answer leaves certain issues undiscussed, three of which will be addressed in the subsequent sections. First, in Section 1.2.2 I discuss the ontological nature of genuine parts. I call attention to the fact that biological systems commonly do not only consist of spatially located objects, but of temporally extended interactions or activities, too. However, the position I endorse is quite moderate (in the sense of being agnostic) since I do not commit myself to the dualistic assumption of Machamer, et al. (2000) that entities (i.e. objects) and activities are two distinct ontological kinds. Second, in Section 1.2.3 I address the vexed question of whether the identification of the genuine parts of a particular system depends on the "theoretical perspective" or "partitioning frame" that a researcher adopts. I argue that even if this is true, it does not imply a strong pluralism, according to which different theoretical perspectives or partitioning frames give rise to different, equally adequate, non-coincident, and (partly) inconsistent decompositions of the same system. Third, in Section 1.2.4 I outline different kinds of decomposability, which point to different types of biological systems (namely to aggregative, component, integrated, and non-decomposable systems). The latter issue is particularly relevant for the third characteristic of reductive explanations that I present in Section 4.

1.2.1 Different Views about Parthood

An earthworm might be decomposed into parts by cutting it (mentally, of course) into one centimeter thick slices. In contrast, a biologist would prefer to dissect the earthworm for example into its segments (which can be more than one hundred), or into its digestive system, its circulatory system, and its nervous system. Are there any reasons not to refer to the one centimeter thick slices as "parts" of an earthworm? Generally speaking, how do we individuate the parts of a system? That is, how do we specify part-whole or compositional relations? Is everything that is located inside of a system already a part of the system? Or are there grounds on which we can say that the one centimeter thick slices are not "proper" parts of an earthworm?

I propose to distinguish three different notions of parts. First, according to classical extensional mereology (CEM)[15] the one centimeter thick slices count as much as parts of an earthworm as its segments or as its digestive, circulatory, and nervous system. Let us call these the *mereological parts* of a system. On CEM, a particular decomposition of a system S is a set of non-overlapping entities such that any part of S overlaps some of them. The most obvious formal properties of the part-whole relation are its transitivity and asymmetry. That is, the following principles are true for any individuals (Simons 1987, 10):

(i) If one thing is a part of another, then the second is not a part of the first.
(ii) If one thing is a part of another, and the second is a part of a third, then the first is a part of the third.[16]

These few details about CEM suffice for my purposes. What should have become clear is that CEM provides us with a quite *unrestricted* notion of a mereological part. According to this notion not only the fourteenth segment or the nervous system count as parts of the earthworm, but also a one centimeter thick slice of it as well as its anterior half.

In biology it seems as if the notion of a part is used in a more restricted way. Commonly, biologists do not claim that an earthworm is composed of one centimeter thick slices. That is, they would refrain from characterizing a one centimeter thick slice as a part of the earthworm – at least, they would not refer to it as a *genuine part* of the earthworm.[17] The underlying view is that nature cannot be arbitrarily divided into certain parts, but that we should in a way "carve nature at its joints".[18] This implies identifying the genuine parts of a system, which constitute a subset of the set of all mereological parts. As I understand the term, genuine parts are the result of a *privileged* decomposition of a biological system. In

---

[15] In fact, there are different formulations of this theory (see Simons 1987, Chapter 1 and 2).

[16] According to CEM, the relation 'x is a part of S' even includes cases in which x and S are identical – even if in these cases x is called an "improper part" (Simons 1987, 11) of S. However, this is not the main respect in which the notion of a mereological part is quite unrestricted.

[17] Note that my usage of the term 'genuine part' deviates from the concept of a proper part. On CEM, what distinguishes proper parts from improper parts is that nothing can be a *proper part* of itself, whereas something can be a part of itself (Simons 1987, 10f).

[18] This does not imply a kind of monism, which holds that there is just *one* right way to carve nature at its joints.

other words, genuine parts of systems are individuated according to certain *principles*. By contrast, non-genuine parts are disclosed by dissecting a system arbitrarily.

This gives rise to the question of what these principles are that guide the decomposition of a system into its genuine parts. Johnston characterizes them as "principles of unity" (2006, 652) and states that these are the principles that bind together the parts into a whole. This goes nicely with the characterization of biological systems I presented in Section 1.1. There I argued that only those entities and interactions are parts of a system that, first, contribute to bringing about the typical behaviors/functions of this system and, second, that are contained within the spatial boundary of this system. This characterization reveals that spatial considerations are insufficient to distinguish the genuine parts of a biological system. What binds together the parts of a system into a unified whole is not only the fact that they share the same spatio-temporal location (i.e. that they are all located inside a particular spatial boundary). What is more, all genuine parts of a system are necessary for the functioning of the system. Hence, what distinguishes the subset of genuine parts from broader set of mereological parts of a system is that only genuine parts contribute to the system's function. For example, the digestive system, the circulatory system, and the nervous system are all genuine parts of the earthworm because they contribute to the growth, reproduction, and survival of the earthworm. The same cannot be said about the one centimeter thick slices of the earthworm.[19]

By emphasizing that genuine parts must be identified with respect to all behaviors/functions that a system typically displays I do not deny Kauffmann's central thesis that systems are decomposed with respect to *one particular* behavior of a system at a time (1970, 259f), rather than with respect to all crucial behaviors at once. However, I maintain that what counts as a genuine part of a system and what does not is, in the majority of cases, not determined on the basis of just *one* behavior of a system, but rather of *several* behaviors – namely on the basis of all those behaviors that are typical for a system (see also Section 1.1.1). For instance, the genuine parts of a paramecium are individuated with respect to the special kind of movement of the paramecium, with respect to its reproduction, with respect to the way it gathers and digests food, and so on. Focusing merely on one of these behaviors would yield only a subset of all the genuine parts a paramecium is in fact composed of.

These considerations lead to a third view about parthood, which has already been touched upon in Section 1.1.1. In the literature on biological mechanisms and mechanistic

---

[19] One can plausibly argue that genuine parts of systems should meet additional restrictions, too. Craver provides an extensive list of what these additional criteria might be. He argues that parts should exhibit four features: first, they should have a "stable cluster of properties" (2007a, 131) and be "loci of stable generalizations" (2007a, 190); second, they should be robust, that is, "detectable with a variety of causally and theoretically independent devices" (2007a, 132; see also Glennan 2002, 344 and Wimsatt 1981); third, one should be able to use them for interventional purposes; finally, they should be "physiologically plausible" (2007a, 132). One can dispute whether the third and fourth feature are necessary conditions that must hold for all kinds of parts. But despite this, these requirements seem appropriate.

explanations a third notion of a part can be identified. Craver offers the most extensive account of what the parts of a mechanism are. He distinguishes what he calls the "*components*" of a mechanism from mere "pieces" by pointing out that only the former "make identifiable contributions to the behavior of a mechanism" (2007a, 187f).[20] In particular, Craver spells out the notion of constitutive relevance by the notion of mutual manipulability (2007a, 139-159; 2007b). In his view, x is a component of a certain mechanism M if there is some change in the behavior of x that changes the behavior produced by M and if there is some change in the behavior of the mechanism M that changes the behavior of x. Briefly speaking, the behavior of a mechanism and the behavior of one of its components must be mutually manipulable. In sum, the components of a mechanism are identified in a similar way as the genuine parts of a system are individuated. However, there exist two important differences: first, mechanisms are always mechanisms for *one particular behavior* (see Section 1.1.1). Accordingly, they are composed exclusively of those entities and interactions or activities that are relevant to this *single* behavior. As opposed to this, the genuine parts of a system include *all* interacting entities that contribute at least to one of the several behaviors that are characteristic for a system. Second, I argued that genuine parts of a system must be located inside a spatial boundary of the system and that these boundaries frequently fall together with certain boundaries that are given in nature, such as membranes that divide an organism into compartments or separate it from its environment (see Section 1.1.3). With regard to the boundaries of mechanisms, Craver explicitly rejects this assumption. He states that "mechanisms frequently transgress compartmental boundaries" (2007a, 141). According to Craver, the *only* criterion that determines whether an interacting entity is a component of a particular mechanism is whether it is *relevant* to the behavior of this mechanism.

Let me summarize. In this section I distinguished three different notions of parts:

(1) *mereological parts* (i.e. everything that counts as a part of a system according to CEM)
(2) *genuine parts* (i.e. any interacting entity that, first, contributes to bringing about at least one of the behaviors/functions a system typically exhibits and that, second, is located inside the spatial boundary of the system)
(3) *components* (i.e. any interacting entity that is relevant to the behavior of a particular mechanism)

(2) is clearly a subset of (1). (3) might be a subset of (2), but in several cases the set of all components of a mechanism will also include entities that are not genuine parts of the corresponding system (more on this in Sections 1.3.3 and 1.3.4). All in all, I think the notion of a genuine part most adequately captures how biologists talk about and identify parts of

---

[20] In the context of explanation this means that only descriptions of components are constitutively explanatorily relevant (2007a, 140).

biological systems, such as the earthworm's segments or its digestive system, circulatory system, and nervous system. The notion of a mereological part is too unrestricted since it provides no means to exclude one centimeter thick slices as parts of the earthworm. Craver might be right that the notion of a component captures how biologists identify the parts of a *mechanism*. However, when it comes to the parts of biological *systems* the mechanistic notion of a component is deficient because it fails to account for two points: first, the parts of systems are commonly individuated not with respect to *one* behavior of a system, but with respect to *all* behaviors that a system typically displays (see Section 1.1.1). Second, in most cases systems are surrounded by a spatial boundary and do *not* transgress compartmental boundaries. This is why factors which are located outside this boundary do not count as parts of the system in question, even if they are relevant to bringing about a certain behavior of the system (and thus would be characterized as a component). Hence, in the following the term 'part' is understood as referring to *genuine parts* of biological systems. Let us now turn to the issues concerning parts and decomposition that are not approached by characterizing parts as genuine parts.

1.2.2 The Ontological Nature of Parts

It is time to prevent a misunderstanding. In the preceding sections I repeatedly claimed that the parts of a system are located inside the spatial boundary of this system. I referred to these parts as entities or as interacting entities. In the examples I discussed the question was often whether certain objects belong to particular systems, for instance, whether centrioles are parts of a cell or whether the house construction is a part of the Caddisfly larva. This focus suggests that *objects* are the primary or only parts of biological systems, and that questions about their *spatial localization* have priority. This impression is reinforced by the work of philosophers who analyze notions of parthood, composition, and constitution (e.g. Evnine 2011; Johnston 1992, 2006; Simons 1987).

However, Nagel already made clear that the story is not that simple. In his paper on "Wholes, Sums, and Organic Unities" (1952), which is republished in his "The Structure of Science" (1961, 380-397), he lists different senses in which the terms 'whole' and 'part' is employed (1952, 18f):

(1) The whole has a *spatial extension* and the part is spatially included in it.
    a. The whole is some length, area, or volume, which contains as parts lengths, areas, and volumes.
    b. The whole is a spatially extended thing with a non-spatial property and the parts are spatial parts with an identical property.
(2) The whole is a *temporal period* and the parts are temporal intervals in it.
(3) The whole is a *class, set, or aggregate of elements* and the parts are subclasses/ subsets or elements of it.

(4) The whole is a *property* of an object or process and the part is some analogous property, which stands in certain relations to the first.
(5) The whole is a *pattern of relations* between certain kinds of objects or events and the parts are either certain elements or classes of elements of this pattern or subordinate patterns.
(6) The whole is a *process* and its parts are subprocesses.
(7) The whole is a *concrete object* and the parts are its properties.
(8) The whole is a *system* and the parts are spatial parts, which stand to each other in various relations of dynamical dependence. More specifically, the parts of systems are either spatially extended constituents or properties/states of the system/its spatial parts or processes or the spatial and dynamical organization.

Nagel's primary goal was to reveal the ambiguity of the terms 'whole' and 'part', not to provide an extensive classification. Accordingly, he does not mention the important fact that most of the above options are *not mutually exclusive*. This fact is only hinted at in the last option and in Nagel's subsequent remarks on organic wholes. There he speaks simultaneously about parts as being "part-processes" (1952, 26) and "constituents" (1952, 27). Moreover, he points out that systems consist of "constituent parts and processes" that "stand to each other in relations of mutual causal interdependence" (1952, 27). In other words, the parts of a system display a specific organization and they are "dynamically interrelated" (1952, 30).

Nagel's way of speaking discloses a feature that is representative for many discussions about systems and their parts. He uses the term 'part' in a broad and in a narrow sense. According to the *broad sense* everything that belongs to or composes a system is called 'part'. This includes all the different ontological kinds that Nagel lists: objects, properties, patterns (i.e. relations), processes, and so on. In contrast, the *narrow sense* of 'part' refers only to objects. This narrow understanding gives rise to expressions such as "constituent parts and processes" (1952, 27), where 'constituent parts' refers to objects and, thus, must be distinguished from constituent processes (which are only parts according to the broader sense of 'part').

However, the narrow notion of a part carries with it the danger that, if it is used primarily or exclusively, it obscures the crucial fact that biological systems are not mere sets of objects. Rather, the objects that compose biological systems exhibit specific properties and relations to each other. More importantly, these relations are not static, but *change* over time. In other words, the objects that compose a system must undergo changes in order to bring about a certain behavior of the system. Consider the example of the transformation of light energy into chemical energy in plants, which is also called the light reaction of photosynthesis. This behavior of plant cells is the result of a whole sequence of changes of certain objects in and around the thylakoid membrane of a chloroplast. Photons are absorbed by photosystems, electrons are transported down an electron transport chain, the released energy is used to pump hydrogen ions across the thylakoid membrane, the created electrochemical proton gradient is employed to generate ATP, etc.

Examples like these led some mechanists to claim that the changes of or interactions between objects constitute a distinct ontological category, the "activities" (Machamer, et al. 2000, 3; Machamer 2004). According to their dualistic view, a biological system is decomposed into *entities* (i.e. objects), exhibiting certain properties, and into *activities*, which they describe as the "producers of change" (2000, 4) and which cannot be reframed as interactions (i.e. as "occasion[s] on which a change in a property of one part brings about a change in a property of another part"; Glennan 2002, 344). As I have pointed out in Chapter II, Section 3.2.2 I do not want to take a side in the dispute about whether monism or dualism is the correct ontology of mechanisms (see Machamer, et al. 2000; Tabery 2004; Torres 2008). Neither am I inclined to interfere in a similar discussion that could be raised with respect to the parts of biological systems. In my view, biological practice does not provide us with sufficient reasons for endorsing either monism or dualism. Rather it supports both views. Hence, we need different kinds of arguments and criteria of adequacy (see Chapter I, Section 2) for choosing between these two metaphysical positions. But this discussion lies outside the scope of this book.

However, in order to make my point I do not need to take a stance in this metaphysical dispute. The reason is that monists as well as dualists agree on an important point. Both state that biological systems encompass not only objects, but also causal interactions between these objects, respectively, activities that the objects engage in. These interactions/activities are the reason why the properties of the constituent objects of a system change over time (which is why these systems are called dynamical). Accordingly, a system not only has a characteristic spatial extension, but involves also typical temporally extended interactions between/activities of its constituent objects. One may argue that a dynamic system consists of particular *processes*, which together form one or more processes that bring about the system's behavior(s). These part-processes are not only spatially, but also *temporally* located within the system. In accordance with this, Bechtel identifies two types of decomposition: *functional* decomposition, that is, the identification of the operations or functions that the behavior of a system can be divided into (which can also be described as interactions or activities), and *structural* decomposition, that is the identification of the spatial parts (i.e. objects) of the system that perform these operations (2006, 31; 2008, 14).[21]

To conclude, if we want to determine the ontological nature of genuine parts of biological systems we cannot simply choose one of the first seven options that Nagel lists and discard all the others (this is why Nagel conflates several of them under (8)). Biological systems consist not only of spatially located objects, but of temporally extended interactions or activities, too. However, as I have already emphasized, one can acknowledge this fact

---

[21] For reasons of simplicity this point is not always sufficiently exposed, which is why I emphasize it here.

without being committed to the dualistic assumption that entities (i.e. objects) and activities (or operations) are two distinct ontological kinds.

1.2.3 Do Parts Exist Prior to Investigation?

The goal of this section is to address the much discussed question of whether the identification of the genuine parts of a particular system is relative to the "theoretical perspective" or "partitioning frame" that a researcher adopts, and whether this suggests a strong pluralism, according to which different theoretical perspectives or partitioning frames give rise to different, equally adequate, non-coincident, and (partly) inconsistent decompositions of the same system.

Another way to frame these questions is to ask whether genuine parts do exist prior to investigation or whether they obtain their "existence" through investigating them. Two authors who explicitly deny the existence of parts prior to investigation are Levins and Lewontin (1985, 2007). They claim that "a whole is a relation of heterogeneous parts that have no prior existence *as parts*" (1985, 273). On their view, not only do "[p]art[s] *make* whole[s]", but also "whole[s] *make* part[s]" (1985, 272). How shall we understand this somewhat peculiar claim? And what reasons can be adduced in favor of such a view? What Levins and Lewontin do not (or, at least, should not) claim is that the entities that are distinguished as being parts of a certain system come into existence through denoting them as parts. This would be a too strong claim because reality as such of course pre-exists. Rather, they seem to assert that the classification of certain entities *as parts* of a system depends on investigation. Let us have a closer look on how Levins and Lewontin argue for this assumption. They state that a system is not composed of "fixed natural units" because "the 'correct' division of the whole into parts varies, depending upon the particular aspect of the whole that is in question" (1985, 272). A similar view can already be found in the work of Kauffmann:

> A view of what a system is doing… supplies criteria by which to decide whether or not a proposed portion of the system with some of its causal consequences is to count as a part and process of the system. […] [D]istinct views of what an organism [or a system in general] is doing may lead us to decompose it in distinct ways. (1970, 259f)

I have already touched on this thesis in the preceding sections. Kauffmann persuasively argues that, depending on which behavior of a system is of concern, the decomposition of that system varies, that is, it discloses different parts. This must not be the case, but often it is the case. The reason is that, mostly, different entities and interactions are relevant for bringing about different behaviors of a system (and only the relevant interacting entities are individuated as parts). For instance, various DNA- and RNA-molecules, ribosomes, and free amino acids contribute to the synthesis of proteins in a cell, whereas other interacting entities, such as centrioles, chromosomes, and the cytoskeleton, are relevant to cell division. Alternatively, consider a particular ecosystem. If the flow of energy in an ecosystem is of interest it typically is decomposed differently (i.e. into different species, environmental factors, and interactions) than if the cycling of nitrogen is of concern.

Understood in this way, I agree with Levin's and Lewontin's thesis that parts do not exist prior to investigation (which means that parts do not exist *as parts* prior to identifying a particular behavior of interest). At this point I should emphasize that this thesis is compatible with my claim that the boundaries of biological systems and their genuine parts are individuated with respect to *all* behaviors that a system typically displays. Ideally, the different decompositions (each of which is made with respect to one particular behavior) can be brought together to form a single view of what the boundaries and genuine parts of a certain biological system are. Cells are not composed *either* of ribosomes, free amino acids, DNA- and RNA-molecules (which are relevant to protein synthesis) *or* of centrioles, chromosomes, and the cytoskeleton (which are relevant to cell division). Instead, they are composed of *all* of these entities (and of many other entities, too), even if they contribute to bringing about *different* behaviors of the cell. However, in some cases such a single view of what the genuine parts of a biological system are might not be more than the wish of a philosopher, which is unachievable in practice. This kind of skepticism can already be found in the work of Kauffmann (1970), but it is particularly put forward by Wimsatt (1974, 2007) and to a certain degree also by Winther (2006, 2011).

Wimsatt argues that different "theoretical perspectives T_i" imply or suggest different criteria for individuating parts, and thus generate different "decompositions $K(T)_i$" (1974, 70). Wimsatt emphasizes that frequently these different decompositions individuate parts whose boundaries are *not coincident*. In such cases there is "*descriptive complexity*" (1974, 70; 2007, 182). The following figure illustrates Wimsatt's thesis:

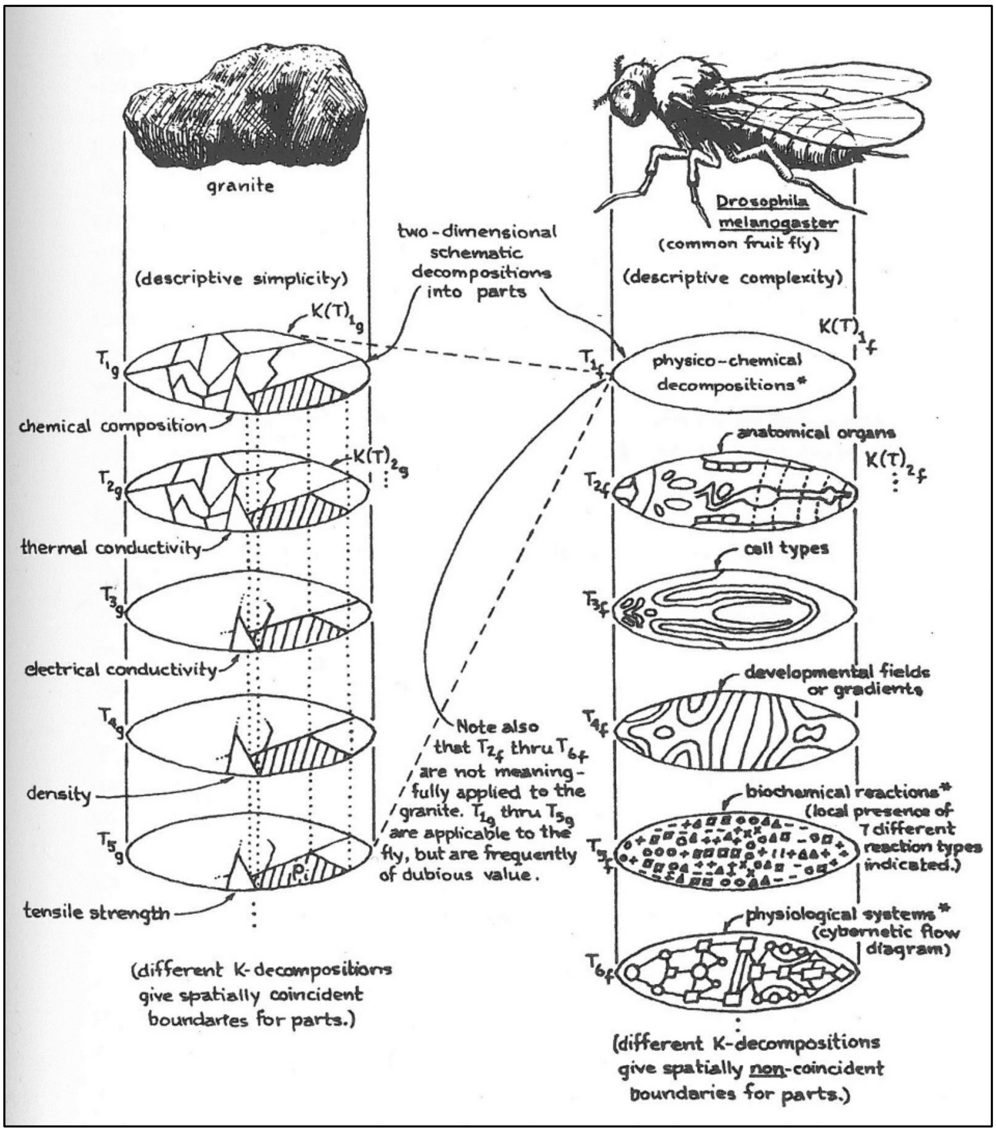

Figure V.6: Descriptive Simplicity and Complexity (Wimsatt 1974, 71)

According to Wimsatt, *Drosophila melanogaster* is a descriptively complex biological system (more precisely, organism) because there are different theoretical perspectives (T_{1f} – T_{6f}) available, which all dissect the fruit fly into different parts, whose boundaries are neither coincident nor isomorphic.

Similarly, Winther recently claimed that in scientific practice there exist certain "theoretical units" that guide the decomposition of a system into its parts and processes (2006, 476). Winther calls these theoretical units "*partitioning frames*" (2006, 475; 2011, 401; my emphasis). He characterizes them as follows:

> A partitioning frame is a set of theoretical and experimental commitments to a particular way of abstracting kinds of parts. For instance, in biology a morphological partitioning frame is distinct from a physiological frame… The former focuses on static, adult structures; the latter identifies processes throughout development. (2011, 401)

According to Winther, partitioning frames include standards for individuating parts and they specify relevant types of part relations (e.g. interaction relations, level relations, and organizational relations). In accordance with Wimsatt, Winther draws attention to the fact that the application of different partitioning frames to the same system gives rise to decompositions, which often fail to map neatly onto one another. However, the picture he paints is not as pluralistic as Wimsatt's. Winther admits that "various sorts of overlap [of different decompositions] are also possible" (2011, 401).

Although Wimsatt and Winther give prominence to Kauffmann's ideas (1970) I take them to defend a thesis that is in a certain respect stronger than Kauffmann's. In a different respect Winther proposes a more moderate view than Kauffmann's and Wimsatt's. Let me elaborate. Wimsatt's and Winther's thesis is stronger than Kauffmann's because they do not merely claim that decomposition depends on the behavior of the system that is of interest. Rather, they argue that decomposition depends on entire theoretical perspectives or partitioning frames. Unfortunately, Wimsatt and Winther do not dwell on this point, but from their few remarks I infer that these perspectives or frames provide *more constraints* on the decomposition of a system than just a specification of the behavior of interest. For instance, they contain several standards for individuating parts and identifying relevant part-relations. At the same time Winther's view seems to be less radical than Kauffmann's and Wimsatt's because he admits that there may be various kinds of overlapping between different decompositions based on different partitioning frames. By contrast, especially Wimsatt stresses the missing conformity between different decompositions.

I share Winther's optimism and reject Kauffmann's and Wimsatt's pluralistic view, according to which different theoretical perspectives give rise to different, equally adequate, non-coincident, and inconsistent decompositions of the same system. Even if there are only a few overlappings between different decompositions of the same system, this does not imply that these decompositions are inconsistent. Wimsatt's figure shows that the parts individuated by different theoretical perspectives may be related to each other via compositional relations. For example, the physico-chemical decomposition of the organism *Drosophila* identifies particular molecules as genuine parts of the organism, which themselves are parts of other genuine parts (which are identified in other decompositions), such as cells, organs, and developmental gradients. Thus, the different decompositions of the same system that result from different theoretical perspectives can be integrated, for instance by identifying compositional relations between the different genuine parts. To be clear, I do not want to endorse the strong thesis that *all* different decompositions of a system can always be integrated into a *single* view. However, I think that more integration is possible than Wimsatt and Kauffmann suggest.[22]

---

[22] This view is substantiated by the fact that it is possible to identify standards for individuating parts that apply to all scientific disciplines, perspectives, and frames. Examples for such criteria are: stability, robustness, intervention, etc. (see Craver 2007a, 128-133; Glennan 2002; Wimsatt 1981; Winther 2011).

1.2.4 Different Kinds of Decomposability

In this section, I examine a question that is of particular importance to the reductive character of explanations. Namely, what are the conditions under which a biological system is nearly decomposable, minimally decomposable, and when does decomposition fail altogether? The answer to this question is crucial for my account of explanatory reduction because the failure of decomposition is an important reason for why an explanation of the behavior of a biological system exhibits a non-reductive character (or why explanation itself fails). Moreover, the third feature of reductive explanations I identify (namely that reductive explanations appeal only to parts in isolation; see Section 4) can be spelled out by referring to the difference between nearly decomposable and minimally decomposable systems. The details of what this means will become more apparent in the following sections. In this section I merely introduce those concepts and distinctions that are necessary to understand my theses.

In their book on "Discovering Complexity – Decomposition and Localization as Strategies in Scientific Research" (2010) Bechtel and Richardson identify different *types of biological systems* by distinguishing different degrees or kinds of decomposability. Here is an overview of the concepts they introduce:

(1) aggregative systems → *strict/simple decomposability*

(2) composite systems
    a. component systems → *near decomposability*
    b. integrated systems → *minimal decomposability*

(3) non-decomposable systems → *failure of decomposition*

The most extreme kind of a decomposable biological system is an *aggregative system*. Wimsatt has developed the most extensive analysis of the notion of aggregativity (e.g. 1986; 1997; 2006b; 2007, Chapter 12). He specifies four conditions that seem separately necessary and jointly sufficient for aggregativity (1986, 260-269). Each of these conditions requires that the behavior(s) of a system must remain invariant to certain modifications of the system (respectively, the system's parts).

(i) *Intersubstitution*: Invariance of the system's behavior under interchanging its parts with parts from a relevant equivalence class.
(ii) *Size scaling*: Qualitative similarity of the system's behavior under changes in the number of parts.
(iii) *Decomposition and Reaggregation*: Invariance of the system's behavior under decomposing and rearranging its parts.
(iv) *Linearity*: Not cooperative or inhibitory interactions among the parts.

Bechtel and Richardson place special emphasis on the last condition since they are primarily concerned with exploring variations in organizational structure. However, all three concur that "[f]ew interesting dynamic systems are strictly aggregative" (Bechtel/Richardson 2010, 25).[23] As soon as the systemic properties are at least partially determined by the organization of the system, we no longer have aggregativity. Consider, for instance, the Na^+/K^+-ATPase, an ATP-driven antiporter that actively transports Na^+ out of the cell and pumps K^+ in against their electrochemical gradients. This transmembrane protein consists of several subunits that interact with each other in a specific way to bring about the behavior of concern (i.e. the antiport of Na^+ and K^+). The following scheme illustrates this:

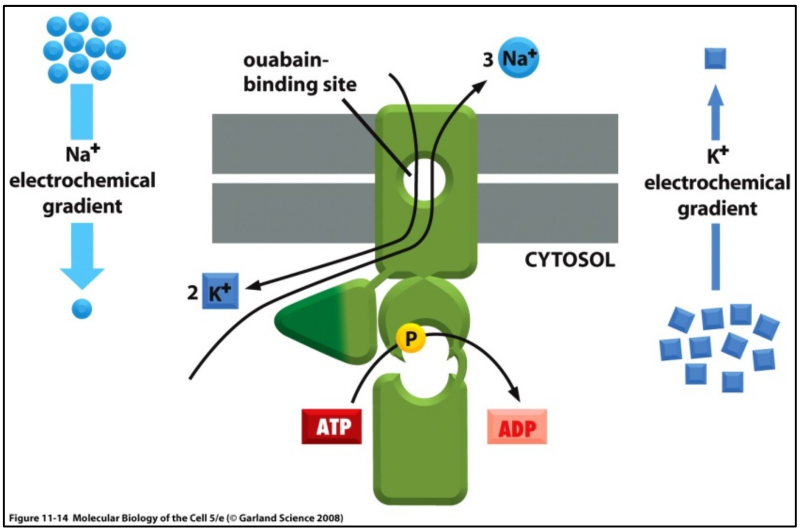

Figure V.7: The Na^+/K^+-ATPase (Alberts, et al. 2008, 662)

The Na^+/K^+-ATPase is far from being an aggregative system because it matters to the behavior of the system what its parts are and how they interact with each other. The parts of the system might be intersubstitutable to a limited degree (since polypeptide chains with varying amino acid sequences may fulfill the same function; condition (i)), but their number cannot be varied (condition (ii)), the parts cannot be rearranged (condition (iii)), and most importantly, the interactions between the parts cannot be neglected (condition (iv)). Hence, the Na^+/K^+-ATPase is what Bechtel and Richardson call a *composite system*.

There are two subtypes of composite systems, namely component systems and integrated systems. The respect in which these two kinds of biological systems differ is the role that is played by systemic organization. Bechtel and Richardson argue that in *component systems* the properties of the parts are "intrinsically determined" (2010, 26). In other words, the causal interactions within the parts are more important for determining their properties than the interactions that occur between parts. As a consequence, it is feasible to determine the properties of the parts *in isolation* (what exactly this means will be examined in Section 4). Component systems are what Simon has called in his influential

---

[23] Similarly, Wimsatt admits that "[i]t is rare indeed that all of these conditions are met" (1997, 375).

paper on “The Architecture of Complexity” “*nearly decomposable*” (1962, 473; my emphasis). By this he means that in these kinds of systems “intra-component linkages are generally stronger than intercomponent linkages” (1962, 477).[24] However, it is not the case that *no* interactions take place between the parts of component systems. Interaction and organization are just not critical since they provide only secondary constraints on the functioning of constituents (2010, 26, 171). In nearly decomposable systems the contributions that the parts make to the behavior of the system are sequential and linear, which is why the components “retain an integrity of their own” (2010, 199). On the basis of this characterization the question emerges whether the Na$^+$/K$^+$-ATPase can be characterized as a nearly decomposable component system. Is it possible to decompose the Na$^+$/K$^+$-ATPase into parts, whose (functional) properties can be discovered “in isolation” (see Section 4) and which contribute sequentially to the behavior as a whole. At first sight, this question can be answered affirmatively. As the following figure illustrates, each of the genuine parts of the Na$^+$/K$^+$-ATPase (e.g. its different subunits, ATP, Na$^+$, and K$^+$) performs a specific task that serves as the input to the operation at the next stage. That is, the interactions can be said to be sequential. For instance, a particular subunit binds and phosphorylates ATP, which induces a conformational change in the protein. This, in turn, induces that another subunit releases Na$^+$ on the outside of the cell.

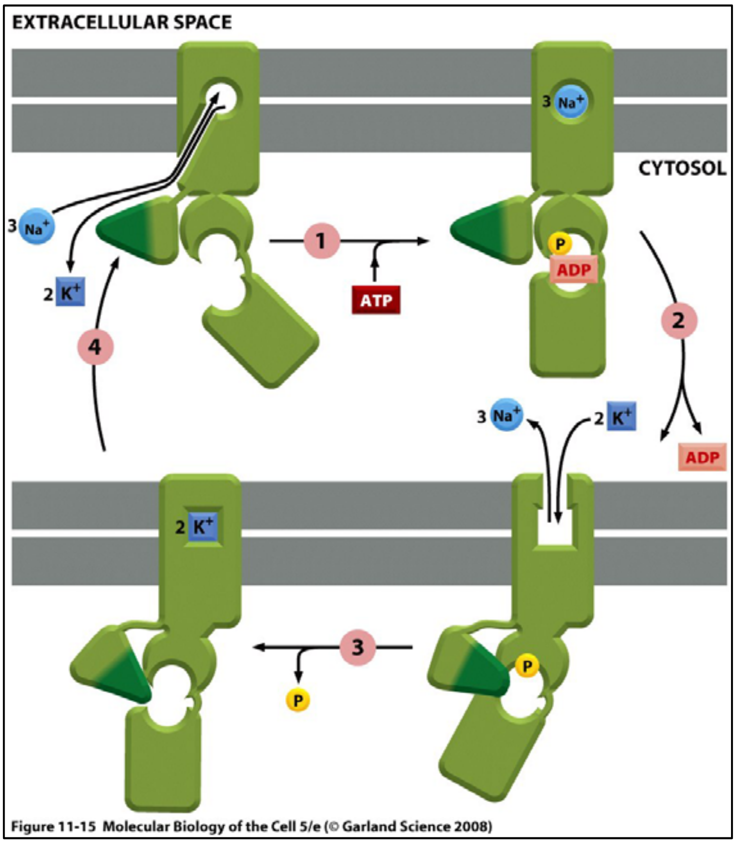

Figure 11-15 Molecular Biology of the Cell 5/e (© Garland Science 2008)

Figure V.8: The Working of the Na$^+$/K$^+$-ATPase (Alberts, et al. 2008, 662)

---

[24] See also Simon 1973. Wimsatt calls this kind of systems “interactionally simple” (1974, 72; 2007, 184).

The properties of the genuine parts of the Na^+/K^+-ATPase can be said to be intrinsically determined. For example, it can be argued that the dispositional property of a particular subunit to bind Na^+ or K^+ (when it is present) is due to the properties of the amino acid residues it is composed of. Accordingly, the behavior of the parts of the system can be studied in isolation.

By contrast, in (functionally) *integrated systems* the properties of the parts are no longer intrinsically determined. Rather, systemic organization is said to be significantly involved in determining constituent functions (2010, 26). This is the case because systems of this kind exhibit a more complex form of organization (such as feedback relations and cyclic organization) and various interactions between the parts. This makes the behavior of the parts "mutually interdependent" (2010, 170) and restricts their autonomy.[25] According to Bechtel and Richardson, the components of integrated systems have no "independent, isolable function" (2010, 31) and are thus only *minimally decomposable*.[26] Integrated systems are sometimes studied by assuming near decomposability. That is, they are dissected into isolated components, which are assumed to make an independent, linear contribution to the behavior of the system. Bechtel and Richardson underline that such a procedure "may blind us to critical factors governing system behavior; in particular, it may blind us to the importance of systemic interaction" (2010, 32). This issue will be picked up on and further elaborated in Section 4.

Finally, Bechtel and Richardson concede that there exist systems, which are not decomposable into distinct components at all and therefore decomposition appears to be "hopeless" or "misguided" (2010, 202) since the system's parts do not seem to perform intelligible subtasks contributing to the overall behavior of the system.[27] In *non-decomposable systems* the behavior of the system is not determined by the contribution of the parts, but by their organization. Accordingly, systemic properties cannot "be anticipated by focusing on the contributions of component units" (2010, 228).

1.3 Levels of Organization

The idea of reduction is closely connected to the idea that nature (or science) is organized into levels and that reduction relates units from different levels. This is true for all kinds of reduction. Ontological reduction is perceived as a relation between, for example, objects, properties, or processes that are located on a higher level of organization and objects, properties, or processes from a lower level. Theory reduction relates higher-level to lower-level theories. And an important element of my account of explanatory reduction is that in

---

[25] Nagel has already emphasized this point: "What is distinctive of such systems [i.e. of organic wholes]…is that their parts do not act, and do not possess characteristics, *independently* of one another." (1952, 26)

[26] This high degree of integration can be traced back to the fact that the component subsystems have evolved together (Levins 1970, 77).

[27] Bechtel and Richardson primarily discuss "connectionist systems" as examples for non-decomposable systems (2010, 199-229).

the case of reductive explanation the phenomenon to be explained is located on a higher level than the explanatorily relevant factors. This is not all there is for an explanation to be reductive, but it is a significant feature.

The view of nature being divided into a hierarchy of levels of organization[28] is prevailing.[29] In reference to Simon's classical paper on "The Architecture of Complexity" (1962) Wimsatt states that

> levels of organization are a deep, non-arbitrary, and extremely important feature of the ontological architecture of our natural world. (2007, 203)

Moreover, levels are not a philosophical invention. They are ubiquitously recognized in the biological sciences as well. Many biological textbooks start with an illustration of biological levels of organization, ranging from ecosystems, communities, and populations over organisms, organs, and tissues to cells, cell organelles, and molecules (see for instance Reece, et al. 2011, 50f). However, despite its ubiquity the notion of a level is notoriously *ambiguous* and frequently left unanalyzed. Thus, it remains a controversial issue and a currently unresolved problem what it means for an entity to be located on a higher or lower level than another entity. This is not to say that we have no clue how to narrow down what an adequate explication of the concept of a level would be. Moreover, there has been philosophical work done in order to illuminate the notion of a level of organization that life scientists employ. Most notably, Wimsatt (1976b; 2007, Chapter 10) and more recently Craver (2007a, Chapter 5) have proposed instructive accounts of levels. They will be discussed below. Beforehand, let me stress an important point.

The concept of level that I (and most other philosophers) seek is an analysis of what a *level of nature* is, not of what a level of science is. The difference between the two is that levels of science relate epistemic units, such as products of science (e.g. theories or explanations) or units of science (e.g. disciplines or fields); whereas levels of nature relate entities in the world, that is, ontological units, such as objects, properties, or processes. My claim is that in the case of reductive explanation the phenomenon to be explained is located on a higher level of organization than the explanatory relevant factors. This claim presupposes that levels are things in the world (i.e. levels of nature), not levels of science. Before I turn to Wimsatt's and Craver's account of levels of nature, I first examine a traditional view of levels, which closely connects the notion of a level of nature to the notion of a level of science.

---

[28] The term 'levels of organization' indicates that I am concerned with the dynamics and structure of hierarchically organized systems. That is, I focus on systems that "consist of smaller units contained in larger ones, in such a way that the lower-level units provide material for the arrangements at upper levels and the upper level arrangements constrain and thus control the activities of the lower levels" (Grene 1987, 504f).

[29] However, I should mention that there are authors who reject the view that nature is hierarchically divided into levels of organization altogether (e.g. Ladyman/Ross 2007, 55-57).

1.3.1 The Layer-cake View of Science

Oppenheim and Putnam define levels by reference to divisions in science rather than by reference to divisions that can be found in the world itself. In their paper on "Unity of Science as a Working Hypothesis" (1958) they identify six levels of science, each of which exhibits a unique set of explanatory principles (i.e. law statements) that constitute a level-specific theory, and each of which uses a unique theoretical vocabulary. In their view, these levels of science correspond to six levels of nature, namely the levels of elementary particles, of atoms, of molecules, of cells, of organisms, and of societies (1958, 4, 9f). What is not of interest at this point, but should be mentioned yet, is that Oppenheim's and Putnam's layer-cake view of science is strongly reductionist. They claim that each theory can be reduced by (i.e. logically derived from) the theory that is located on a level beneath, which gives rise to the unity of science.

Several philosophers have criticized this layer-cake view of science and the related notion of levels of science for their failure to correspond to actual scientific practice (Waters 2008, 247-249; Craver 2007a, 172-177, to mention only some recent criticism). Briefly speaking, Oppenheim's and Putnam's six-level-image of science and of the world is too simplified. But even if we grant that this view might be enriched with further intermediate levels (e.g. with the level cell organelles, of tissues, of organs, of populations, of ecosystems, and so on) the objection would still apply that scientific theories and in particular scientific disciplines do not correspond to single levels of nature. Instead, most biological theories and disciplines span different levels. That is, they make assertions about objects, properties, and processes that are located on more than one level of organization. Conversely, the objects, properties, and processes, that are located on one single level of nature, are approached by more than one scientific theory and often also by more than one scientific discipline.

To conclude, there is no neat correspondence between the levels of science (i.e. the units or products of science, such as theories and disciplines) and levels of nature, like the one Oppenheim and Putnam (1958) imagine.

1.3.2 Levels as Local Maxima of Regularity and Predictability

The core idea of Wimsatt's approach is that objects can be assigned to the same level if they are of *comparable size* (and, associated with this, exhibit comparable dynamical properties; 1976b, 237). This idea is further refined by the assumption that a level of organization is a "*local maximum of predictability and regularity*" (1976b, 238) in the phase space of possible ways of organizing matter. The following figure illustrates this thesis:

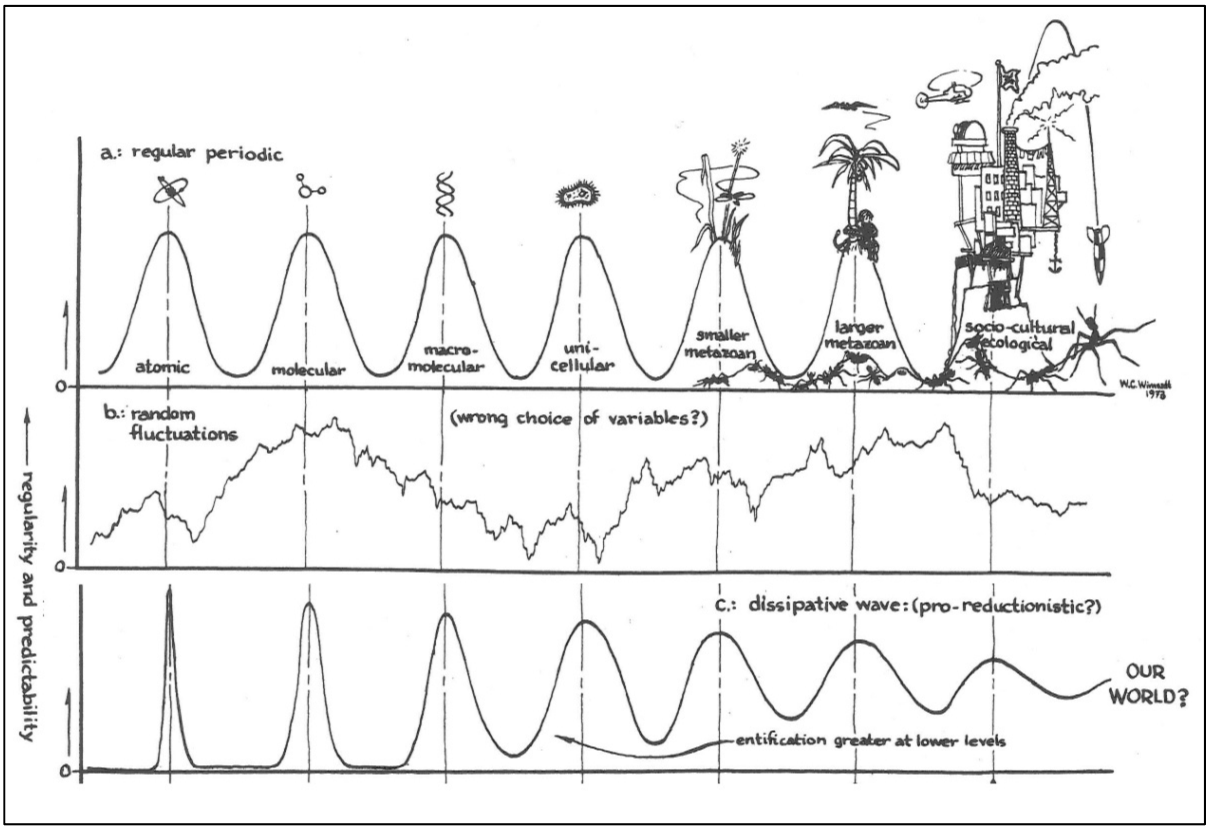

Figure V.9: Levels as Local Maxima of Regularity and Predictability (Wimsatt 2007, 224f)

In this figure levels appear as peaks of regularity and predictability when graphed against a roughly logarithmic size scale. The above graph represents a world that is neatly ordered into levels. The graph in the middle depicts a world in which there is random fluctuation of regularity and predictability, that is, in which there exist no levels of organization. The bottom graph represents what is probably the world in which we actually live. The fact that the amplitude of the waves decreases with rising size illustrates Wimsatt's assumption that "the higher levels seem to become more diffuse – less regular and less localizable" (1976b, 254) – than the lower levels. This is due to the fact that the richness of causal relations within and between levels increases at higher levels of organization, which is why they are not as neatly separable as lower levels are.

At some point the concept of a level even breaks down completely as a useful way of characterizing the structure of the world. At this point "other ontological structures enter, either as additional tools or as replacements" (2007, 222). These other ontological structures are what he calls "perspectives"[30] (2007, 227) and "causal thickets"[31] (2007, 237). What is

---

[30] "In complex systems, *perspectives* give organized approaches to a cluster of problems and techniques, often span levels, cross-cutting levels and each other, and give knowingly incomplete descriptions of the systems to which they are applied. (Levels can be viewed as special cases of perspectives ordered by hierarchical part-whole compositional relations.)" (Wimsatt 2007, 358)

[31] The term *causal thicket* "is intended to indicate a situation of disorder and boundary ambiguities" (Wimsatt 2007, 238).

important to my analysis is that Wimsatt links the breakdown of levels to the absence of "relatively unambiguous inclusion or compositional relations" (2007, 221). In other words, Wimsatt assumes that one can only individuate well-defined levels of organization if one can identify *compositional relations* between the objects assigned to different levels. If these neat compositional relations cannot be found, for instance, because the complexity and context-dependency of a system is high and its modularity and regularity is low, one cannot distinguish levels of organization. Hence, the existence of compositional relations between objects from different levels is – besides size and local maxima of predictability and regularity – another major characteristic of levels.[32] Wimsatt mentions this feature of levels in his early work (e.g. 1976b, 215, 254), but it is particularly emphasized in his newest book:

> By level of organization, I mean here compositional levels – hierarchical divisions of stuff… organized by *part-whole relations*, in which wholes at one level function as parts at the next (and at all higher) levels. (2007, 201; my emphasis)

This quote shows that Wimsatt regards not only size and local maxima of predictability and regularity as important criteria for identifying levels, but also compositional relations. Accordingly, the taxonomy of levels that Craver proposes is instructive, but fails to adequately classify Wimsatt's account of levels.

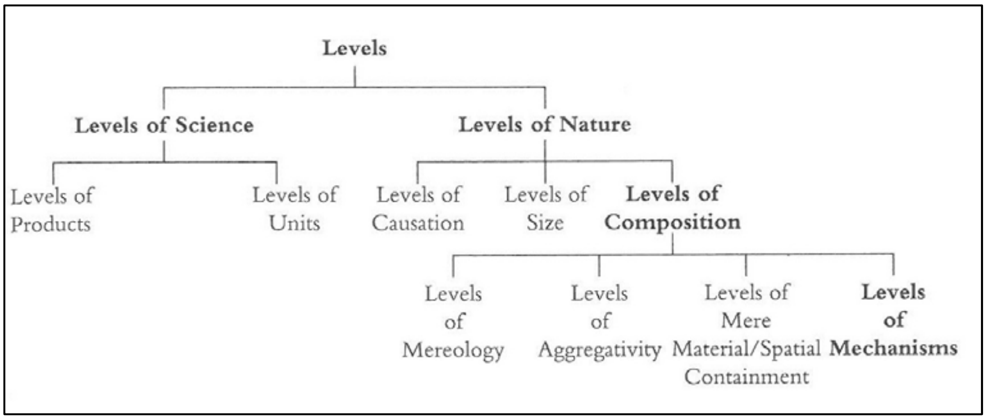

Figure V.10: A Taxonomy of Levels (Craver 2007a, 171)

Craver claims that Wimsatt treats levels as levels of size, whereas he regards levels as being specific levels of composition (namely levels of mechanism; see next section). Such a classification ignores the important fact that Wimsatt proposes not only one, but *different* crucial features of levels. In Wimsatt's view, levels of nature are levels of size, local maxima of predictability and regularity, *and* levels of composition. What is more, although Wimsatt treats size as an important factor in determining levels, he admits that "[s]ize is not a sufficient indicator of level" since there might be objects such as "bacterium-sized black

---

[32] This is why the reductive method of decomposition is also called "downward looking" (Lidicker 1988, 278; Byerly 2003, 337; Bechtel 2009, 543): the parts of a system are located on a level beneath the level of the system. Thus, one looks down a level when one decomposes the system into its parts.

holes" (2007, 208). Thus, Craver is misguided in characterizing Wimsatt's account as an account of levels of size.

However, it must be acknowledged that it is far from easy to pin down what exactly Wimsatt's account of levels is and how the different elements of his account relate to each other. This is due to the fact that Wimsatt refuses to provide an analysis of the concept of a level in terms of necessary and sufficient conditions. Rather he provides us with the following summary:

> They [i.e. levels of organization] are constituted by families of entities usually of comparable size and dynamic properties, which characteristically interact primarily with one another, and which, taken together, give an apparent rough closure over a range of phenomena and regularities. (2007, 204)

He adds a bit scoffingly:

> For all who still believe in 'necessary and sufficient conditions' style analysis, I note at least five qualifiers in this sentence – all apparently necessary – that would be difficult at best to deal with, and the referents of these qualifiers are also often disturbingly general, and correspondingly unclear. […] Definitional language is notoriously unhelpful in contexts like these. (2007, 204)

I agree with Wimsatt that it is sometimes fruitless to analyze a scientific concept by seeking exclusively for necessary and sufficient conditions (and not, for instance, for typical features or generalizations with exceptions). However, this insight should not mislead one to abandon the ideals of clarity and accuracy – two virtues that one sometimes misses when reading Wimsatt's work.

Another objection can be raised with regard to Wimsatt's thesis that levels constitute local maxima of regularity and predictability because objects at the same level "interact most strongly and frequently" (1976b, 215). One could argue that this requirement is too strong because there exist also regular and predictable *interlevel* relations. Let us consider this claim in more detail, at first with respect to causal interlevel relations, thereafter with respect to non-causal interlevel relations (such as compositional relations).

First, the assumption that there are causal interactions between wholes and their parts (intra-systemic causation) is commonly conceived as problematic (see e.g. Craver/Bechtel 2007). However, one could still claim that there exist causal interactions between objects from different levels that are not directly related via part-whole relations. Consider an example. Enzymes are commonly said to be located on a higher level of organization than amino acids because enzymes are composed of amino acids (and, thus, enzymes are larger than amino acids). Wimsatt would add that another reason for assigning enzymes and amino acids to different levels is that enzymes interact most strongly and frequently with other enzymes and amino acids with other amino acids. One could object to this thesis by pointing to the causal interactions that take place between enzymes and amino acids. However, it would be controversial to claim that an enzyme is causally related to the amino acids of which it is composed (Craver/Bechtel 2007). Instead, a whole (i.e. enzyme) is *compositionally* related to its parts (i.e. amino acids). But this is compatible with the claim that an enzyme can be causally related to other amino acids, that is, to amino acids that do not make up the enzyme. For instance, digestive enzymes break down polypeptide chains into free amino acids. Accordingly, digestive enzymes can be said to causally interact with free amino acids. I see no reason why these interlevel causal interactions are less strong, less regular, or less predictable than those between enzymes or between amino acids. However, Wimsatt may be right to argue that interlevel causal relations (e.g. between enzymes and amino acids) are not as common as intralevel causal relations (e.g. between enzymes and enzymes or amino acids and amino acids).

Second, Wimsatt's thesis that interlevel relations show the highest degree of regularity and predictability might be plausible with respect to causal relations, but it is not convincing with regard to non-causal relations, such as compositional relations. In line with this, Craver argues that "there are many interlevel relations… of regularity and predictability" (2007a, 183), for instance regular dependencies of the behavior of the whole on the behavior of its components. However, Wimsatt could respond to this objection either by emphasizing that his claim applies only to regularities and predictions that concern interactions, that is, causal relations, or by acknowledging that there exist regular and predictable non-causal interlevel relations and yet insisting that most regularities and predictions concern relations between objects on the same level. In other words, Wimsatt could admit that the amplitudes of the waves are smaller than originally expected (see Figure V.9), but that one can nevertheless distinguish discrete waves (i.e. local maxima of predictability and regularity).

1.3.3 Mechanistic Levels

In his book on "Explaining the Brain" (2007a) Craver develops an alternative account of levels. The main goal of his approach is to capture the notion of level that underlies current neuroscientific explanations, such as the explanation of spatial memory (2007a, 165-170). Craver's central thesis is that levels of nature are a special *subtype of compositional levels*, namely mechanistic levels.[33]

Let us start with examining two alternative ways of how one could specify the notion of a compositional level. Craver argues that neither classical mereology, nor the concept of material/spatial containment alone succeeds in yielding an adequate concept of level. Briefly speaking, Craver claims that the formal apparatus of most mereologies includes assumptions that are misleading (and, thus, fails to specify necessary conditions), and that the notion of material/spatial containment alone gives rise to a too unrestricted notion of what a part is

---

[33] The view that levels of nature should be specified as being mechanistic levels can be found in the work of Bechtel, too (2008, 141-148; Bechtel/Hamilton 2007, 411-413).

(that is, it specifies only sufficient conditions). In my view, Craver is right about the former, but not about the latter. Consider mereology first.

Craver persuasively demonstrates that the formal apparatus of most mereologies fails to provide necessary conditions for part-whole relations and, thus, for the existence of levels. He identifies three misleading assumptions (2007a, 184-186). First, the reflexivity theorem, which is included in most mereologies, states that every object is part of itself. This contradicts the understanding of the notion of a part that is common to biological practice. According to this understanding, there are no cases in which y is a part of x and, at the same time, x is identical with y. This intuition is captured in classical extensional mereology (CEM) by distinguishing parts from proper parts. For the latter it holds that nothing is a proper part of itself (Simons 1987, 10; see also Section 1.2.1). Second, many formal mereologies encompass the extensionality theorem, which holds that an object is completely determined by the set of its parts. This theorem is problematic since it neglects the importance of the organization of the parts to the identity of the whole. Biologists would not regard two objects, which share all their parts, but differ in how these parts are organized, as being identical. This is why most biological systems are more than the (mereological) sum of their parts. Third, formal accounts of mereology often apply to both abstract and concrete entities. But only those part-whole relations that exist between concrete entities provide an appropriate basis for specifying the notion of a level of nature. The reason for this is that the condition of material and spatial containment is true only for part-whole relations between concrete entities. For instance, one might argue that the holdings of Peter's Swiss bank account are part of his total wealth. But the holdings are not materially or spatially contained in his wealth – also because it is questionable whether his holdings and his total wealth have a definite spatial location at all.

Accordingly, Craver regards material and spatial containment as a crucial feature of part-whole relations, which give rise to the notion of a compositional level. But yet, he argues that "[l]evels of mere material/spatial containment are too permissive to characterize the nature of LM levels [i.e. levels of spatial memory]" (2007a, 187). In other words, Craver asserts that the relation of material and spatial containment is a sufficient, but not a necessary condition for the existence of the kind of part-whole relations that define levels of organization. In particular, he claims that relations of mere material and spatial containment fail to distinguish between "mere pieces of a system" and the "components" (2007a, 187) of a mechanism.

In order to understand this claim, recall the three notions of parts I introduced in Section 1.2.1. There I distinguished between mereological parts of a system, genuine parts of a system, and components of a mechanism[34]. Any entity that counts as a part of a system according to CEM is a *mereological part*. The concept of a *genuine part* refers to those

---

[34] To a first approximation mechanisms can be regarded as subsets of systems. To illustrate this, if systems were round cheeses without holes, a mechanism would be a Swiss cheese (see Section 1.1.1).

entities that meet two further conditions, namely that, first, contribute to bringing about at least one of the behaviors/functions a system typically exhibits and that, second, are located inside the spatial boundary of the system. *Components*, in turn, are those interacting entities that are relevant to a specific behavior of a particular mechanism.

Craver's central thesis is now that levels of nature are determined by *component-mechanism relations*, rather than by one of the two alternative part-whole relations, namely mereological-part-system relations or genuine-part-system relations. Actually, Craver does not provide a critique against the thesis that genuine-part-system relations are the compositional relations that determine the levels of nature (this is the view I endorse). Rather, he only argues that mereological-part-system relations are inappropriate (2007a, 184-186) and that relations of mere material/spatial entailment are insufficient (2007a, 187-190). However, both criticisms do not apply to my thesis that genuine-part-system relations determine the levels of organization since, according to my view, genuine-part-system relations are more than mere spatial relations. But before I discuss these and other objections to Craver's notion of level consider two additional features of his account.

First, Craver states that, "[i]n levels of mechanisms, the relata are behaving mechanisms at higher levels and their components at lower levels". Moreover, he stresses that these lower-level components should be understood as "acting entities", which are "organized together to form higher-level components"[35] (2007a, 189). Craver argues that his view of levels of mechanisms as "levels of *behaving components*" (2007a, 190; my emphasis) deviates from common views of levels, which conceive levels as being levels of objects (e.g. molecules, cells, organs, organisms, populations, etc). Craver may be right that common notions of levels mistakenly take into account only objects and that this is a crucial shortcoming of these accounts. However, in Section 1.2.2 I argued that one can allow for the importance of interactions, temporal changes, and organization of the parts of a system without being committed to the dualistic ontology of the mechanists (i.e. to the claim that entities and activities constitute distinct ontological kinds).

Second, an important feature of Craver's notion of level is that levels of mechanisms are only defined *locally*, that is, only within a given compositional hierarchy. In other words,

> "[h]ow many levels there are, and which levels are included, are questions to be answered on a case by case basis by discovering which components… are explanatorily relevant for a given phenomenon. They cannot be read off a menu of levels in advance." (2007a, 191)

According to this local notion of a level, it makes no sense to ask, for example, whether a muscle fiber (which is a multinuclear cell) is located on a lower level than the liver (which is an organ) because the muscle fiber is not a component of a mechanism that explains one of

---

[35] Craver takes organization to be an "*interlevel relation* between a mechanism as a whole and its components" (2007a, 189; my emphasis). This sounds peculiar to me since we commonly speak about the organization *of parts*. Accordingly, organization would be a set of relational properties of parts, that is, of lower-level properties, rather than an interlevel relation.

the many behaviors/functions of the liver. Claims about levels can only be made with respect to acting entities that belong to the same hierarchy of mechanisms. In short, "X and S are at the same level of mechanisms only if X and S are components in the same mechanism" (2007a, 192). Craver contrasts his account of local mechanistic levels with the view that the world is divided into *monolithic levels of nature*. Such a monolithic image is for instance suggested by Wimsatt's figure of levels as local maxima of regularity and predictability (see Section 1.3.2, Figure V.9). According to Wimsatt, levels of nature are not defined just locally and restricted to particular component-mechanism relations. Rather, they are *global* levels, which allow us to state that muscle fibers are located on a lower level than livers because the former are cells and the latter are organs and cells can be said to constitute genuine parts of organs.

Let us turn to the shortcomings of Craver's account. In my view, Craver's concept of a mechanistic level is *inappropriate* for specifying the notion of a level that is involved in discussions about reductionism and reductive explanation in at least four respects.

First, the thesis that component-mechanism relations are those compositional relations that determine levels of organization is mistaken since it yields a *too restricted* notion of level. Let me explain why. According to Craver, a mechanism is always a mechanism for one singular phenomenon, and only those interacting entities are components of this mechanism, which contribute to bringing about this particular phenomenon. Thus, claims about levels can only be made with respect to a certain mechanism and its components. For instance, consider a paramecium (i.e. a unicellular ciliate protozoon). A paramecium displays several different behaviors, such as gathering food, movement, digestion, reproduction, etc. One may argue that for each behavior there exists a mechanism that brings about this behavior. For instance, there exists a mechanism for gathering food, which consists of cilia that sweep up food along with some water into the cell mouth. And there exists a mechanism for asexual reproduction, which has the micronucleus, the macronucleus, the cell membrane and other interacting entities as components. If we accept Craver's local notion of mechanistic levels we can only make claims such as that the cilia are located on a lower level than the moving paramecium (because the cilia are a component of the mechanism for movement of the paramecium) and that the macronucleus is located on a lower level than the reproducing paramecium (because the genome is a component of the mechanism for reproduction of the paramecium). However, we *cannot* state that the cilia and the macronucleus are located on the same or on different levels since they are not components of the same mechanism.[36] This is an unacceptable consequence because we commonly make such claims as that the cilia and the macronucleus are parts of the same system (namely the paramecium) and as such are located on the same level of organization. We can avoid this consequence by taking

---

[36] Unless one identifies a mechanism that has cilia and the macronucleus as components.

genuine-part-system relations (and not component-mechanism relations) to be the ones that give rise to distinctions between levels. In line with this, I have argued in Section 1.1.1 and 1.2.1 that biological systems do not only consist of those interacting entities that are components of a mechanism for one specific behavior that the system displays. Instead, systems like the paramecium are composed of *all* interacting entities that contribute to bringing about at least one of the behaviors that are characteristic for the system (i.e. of genuine parts). In sum, relations between systems and their genuine parts seem to be a much better basis for distinguishing levels of organization than relations between mechanisms and their components.

Second, Craver's notion of mechanistic levels relies on a notion of the *boundaries of a mechanism* that is linked too closely to the notion of explanatory relevance (which he specifies as constitutive relevance; see Section 1.2.1 and Chapter IV, Section 1.2.2). The core of the problem is that any factor that counts as constitutively relevant and, as such, is referred to in the explanation is *by definition* a component of the mechanism (2007a, 141-144). Let me elaborate on this point. In their popular paper on "Thinking about Mechanisms" (2000) Machamer, Darden, and Craver argue that so called "set-up" or "enabling conditions" and even factors like temperature and pH, which are commonly characterized as background conditions, "are not inputs into the mechanism but are parts of the mechanism" (2000, 11). However, this characterization in particular of set-up conditions as being components of a mechanism and, as such, as being *internal* to the mechanism contradicts the common way of characterizing them as *environmental* factors. Even worse, Craver's thesis that any constitutively relevant factor that is referred to in a mechanistic explanation is by definition a component of the corresponding mechanism renders it *impossible* that a mechanistic explanation appeals to environmental factors or to factors that are located on the same or on higher levels than the mechanism. Mechanistic explanations refer only to components of a particular mechanism and these are by definition *internal* to the mechanism and located on a *lower level* than the mechanism.[37] This view of explanation is inadequate with respect to several biological explanations, which are said to make explicit reference to contextual or to higher-level factors.[38]

---

[37] Some mechanists explicitly deny this claim and state that descriptions of contextual factors and higher-level factors are included in mechanistic explanations (e.g. Bechtel 2008, 2009). However, in order to do so Bechtel must abandon Craver's notion of constitutive relevance. But Bechtel does not provide us with an alternative notion of explanatory relevance, respectively with an alternative way to determine the boundaries of mechanisms.

[38] One could escape this problem (at least with respect to environmental factors) by claiming that the boundaries of a mechanism exceed the boundaries of the corresponding system and that the latter determines whether a factors counts as internal or external. According to this interpretation, for example the incoming nervous signal that triggers muscle contraction could be viewed as a factor, which is external to the system because it is located outside of the boundary of the muscle fiber, but which nevertheless is a component of the mechanism for muscle contraction and thus internal to this mechanism. However, one consequence is that mechanisms could not be characterized as being subsets of systems any more.

Third, the local character of Craver's concept of a level renders it insufficient for capturing the notion of level that underlies discussions about reductionism and reductive explanation in the biological literature. One can only make sense of the notion of a reductive explanation by, at least partially, relying on a global notion of levels. In particular, the notion of just one, fundamental biological level of molecules (see Section 2.3.1) presupposes such a global concept of level. Furthermore, it matters for the reductive character of an explanation on which level environmental factors (which are no genuine parts of the system of concern) are located. We want to say that an explanation is reductive because it refers only to lower-level factors and although it also cites external factors. A local notion of level, like Craver's, would not allow us to make claims about the level, on which external factors are located. In Section 2 and 3 I will further elaborate and substantiate this claim.

Fourth, the characterization of explanations as exhibiting a *multilevel* character presupposes a concept of level that is not only locally defined. In accordance with other authors (e.g. Mitchell 2009, 109-115; Bechtel 2006, 27; Schaffner 2006, 384-387) Craver himself states that neuroscientific explanations span multiple levels (e.g. 2007a, 9-16). By this he does not only mean that mechanistic explanations involve two levels, namely the level of the mechanism whose behavior is to be explained (explanandum) and the level of the components of that mechanism (explanans). In addition, he seems to claim that the explanans itself refers to interacting entities from different levels. However, the latter thesis relies on a global notion of level since according to the locally defined notion of a mechanistic level all constitutively relevant factors are components of the corresponding mechanism and, thus, all located on the same mechanistic level.

1.3.4 The Notion of Level in My Account of Reduction

After having reviewed Putnam's and Oppenheim's, Wimsatt's, and Craver's account of level, the question arises which insights we have gained so far. In particular, has the previous discussion brought to light a notion of level that is suitable for my analysis of explanatory reduction? The preceding examination of different accounts of levels disclosed four issues:

First, what we seek is an account of *levels of nature*, not of levels of science. Moreover, levels of science (e.g. of theories) are a deceptive indicator for levels of nature since different theories from different disciplines refer to entities at the same level and since often one theory appeals to entities from different levels. Second, *size* is no more than a rough indicator of levels. In fact, size relations are "incidental by-products" (Craver 2007a, 183) of more fundamental compositional relationships among entities from different levels. Third, I agree with Craver and Wimsatt that *compositional relations* determine which levels of organizations exist in nature. However, this leaves open the question what kind of compositional relation is crucial. Relations between mereological parts and systems turn out to be inappropriate. Furthermore, the currently best worked out account of levels, namely Craver's notion of a mechanistic level, is promising, but in the end his specification of part-whole relations as component-mechanism relations is unconvincing, too. I will elaborate on this point below. What is left over is the option that the distinction between different levels of organization can be traced back to compositional relations between systems and their *genuine parts*. Fourth, contrary to what Craver claims I argue that a notion of level that is defined only locally is inadequate. More precisely, the concept of a local level is insufficient for clarifying what makes an explanation reductive. What exactly this means will become clear in the remaining sections of this chapter. Fifth, I do not want to preclude that there may be *additional features* of levels of organization, despite the fact that entities at different levels are related to each other by compositional relations. A candidate for such an additional feature is Wimsatt's characterization of levels of nature as local maxima of regularity and predictability. Although Wimsatt's account in general is too vague to be persuasive it contains insights that are worth to be elaborated further.

All in all, neither Wimsatt nor Craver have proposed the *one* account of levels, which is fully convincing and which is appropriate for my analysis of reductive explanations. However, both accounts contain important insights into the nature of levels. For instance, both authors regard compositional relations to be crucial for the distinction between levels. However, Wimsatt fails to provide a specification of which kind of part-whole relation determines levels, and Craver's specification turns out to be inadequate. I argued that component-mechanism relations fail to give rise to a persuasive distinction among levels of organization. Instead, the compositional relations that determine levels should be specified as genuine-part-system relations. What is more, the notion of level that underlies discussions about reductionism and reductive explanation in the biological literature cannot be captured by a local concept of level (at least not by it alone). Biologists make claims, such as that blood cells are located on the same level than neurons – even if blood cells and neurons are not components in the same mechanism. Moreover, they argue that systems can be explained reductively if the environmental factors that are referred to in the explanation are lower-level factors, too. An adequate analysis of reductive explanation must account for this way of speaking. Hence, it requires a notion of level that cannot merely be spelled out locally. In order to get such a concept of global levels we can either invoke Wimsatt's additional criterion of levels being maxima of regularity and predictability, or we can try to extend the notion of a local level, for instance by bringing together different genuine-part-system distinctions and establishing a hierarchy of different genuine-part-system relations. This is for instance the way that Fazekas and Kertész (2011, 379f) choose and that appears to be promising. However, one should be cautious not to impose a desirable but ill-fitting view on reality since it might be that not all genuine-part-system relations can be related to each other and can be integrated into a neat view of what the levels of organization are (see Section 1.2.3).

1.4 Interim Conclusion

The goal of this section was to clarify those concepts that are of particular importance for specifying the notion of a reductive explanation in biology, namely the concept of a biological system and of its context or environment (Section 1.1), the concept of decomposing a system into its parts (Section 1.2), and the concept of a level of organization (Section 1.3). My analysis yielded the following results:

In Section 1.1 I argued that *biological systems* are composed of a set of interacting entities that form together a unified whole. This means that the parts of a biological system satisfy two conditions: one the one hand, they *work together* to bring about those functions/behaviors that the system characteristically displays (Section 1.1.1), and on the other hand, they are located within a defined zone, which is surrounded by a *spatial boundary*. In Section 1.1.2 I endorsed the view that the spatial boundary, which separates the system from its environment, ideally is a *continuous, constant surface*. However, I admitted that it sometimes might be challenging to separate a biological system from its environment. These difficulties trace back to the fact that in some cases the boundaries of systems are subject to continuous change and that in other cases boundaries are not continuous surfaces, but rather conglomerates of several disconnected boundaries (in these cases the system is more a cluster of islands than a round cheese). I claimed that these and other cases (like symbiotic associations, superorganisms, and extended phenotypes) do not show that biologists fail to perform the task of distinguishing system and environment. They merely reveal that the view of the boundary of a system as one continuous, constant surface sometimes is illusory and that biologists do not always agree on how to draw the line between a system and its environment. Section 1.1.3 showed that this, however, does not imply that the separation of a system from its context is a highly flexible or arbitrary matter. I argued that, instead, the boundaries of many biological systems fall together with certain structures that exist in nature (e.g. with membranes). This is why I characterized the boundaries of biological systems as *bona fide boundaries*, rather than as fiat boundaries.

The goal of Section 1.2 was to specify the notion of a part and the notion of decomposition and to answer central questions that relate to these two concepts. In Section 1.2.1 I distinguished three notions of parts: first, everything that counts as a part of a system according to classical extensional mereology is a *mereological part*; second, any interacting entity that contributes to bringing about at least one of the behaviors/functions a system typically exhibits and that is located inside the spatial boundary of the system is a *genuine part*; third, any interacting entity that is relevant to the behavior of a particular mechanism is a *component*. I concluded that with respect to biological systems the notion of a genuine part is the one that is most adequate. In Section 1.2.2 I argued that one should be cautious not to identify parts with *objects* because biological systems are composed not only of spatially located objects, but of temporally extended interactions or activities, too. In Section 1.2.3 I agreed with Kauffmann that the decomposition of a system depends on the behavior of the system that is of interest. However, I emphasized that this does not imply the pluralistic assumption that different theoretical perspectives or partitioning frames result in different, equally adequate, non-coincident, and (partly) inconsistent decompositions of the same system. By contrast, I adopted a more optimistic view and stated that the different decompositions of a system frequently are consistent and can be integrated via other compositional relations. Finally, in Section 1.2.4 I distinguished different kinds of decomposability, namely simple decomposability, near decomposability, and minimal decomposability. I stated that aggregative systems are *simply decomposable* and that their behavior remains invariant to certain modifications of the system's parts. I characterized component systems are *nearly decomposable*, which means that they are composed of parts whose properties are intrinsically determined and can be investigated in isolation. Organization and interactions between the parts are not crucial to the behavior of a component system. I claimed that, opposed to this, in integrated systems the properties of the parts are co-determined by the systemic organization. This makes the behavior of the parts mutually interdependent, which is why integrated systems are only *minimally decomposable*. I announced that the difference between aggregative, component, and integrated systems will be relevant to my analysis of the major features of reductive explanation (see, in particular, Section 4).

In Section 1.3 I argued that the notion of reduction is closely connected to the picture of nature as being divided into a hierarchy of levels of organization, but that the concept of levels of nature remains notoriously unclear. My analysis of this concept revealed that levels of science (e.g. of theories) are a deceptive indicator for levels of nature since different theories from different disciplines refer to entities at the same level and since often one theory appeals to entities from different levels (Section 1.3.1). In Section 1.3.2 I argued that size can be no more than a rough indicator of levels, too, because size relations are only by-products of *compositional relations*, which are the proper relations that determine levels. However, my central claim was that it must be specified which part-whole relations give rise to different levels of organization. Section 1.3.3 showed that mereological-part-system relations are inadequate, just as Craver's component-mechanism relations. I claimed that, instead, genuine-part-system relations seem to be the ones that give rise to a convincing notion of levels of nature (Section 1.3.4). Finally, I pointed out that the notion of level that underlies discussions about reductionism and reductive explanation in biology cannot be captured by a concept of level according to which levels are only defined locally (i.e. with respect to a particular component-mechanism distinction; Section 1.3.3). Rather, I argued that my analysis needs an account of *global levels*. In Section 1.3.4 I claimed that such a global notion of level may be obtained, for example, by bringing together different genuine-part-system distinctions and establishing a hierarchy of different genuine-part-system relations.

To conclude, the investigation of biological systems and the development of explanations for the behaviors of these systems seem to involve three major tasks: first, the individuation of the system of concern (which includes the separation of the system from its environment), second, the decomposition of that system into its genuine parts, and third, the identification of different levels of organization and the association of certain interacting entities to these levels. The following figure illustrates this:

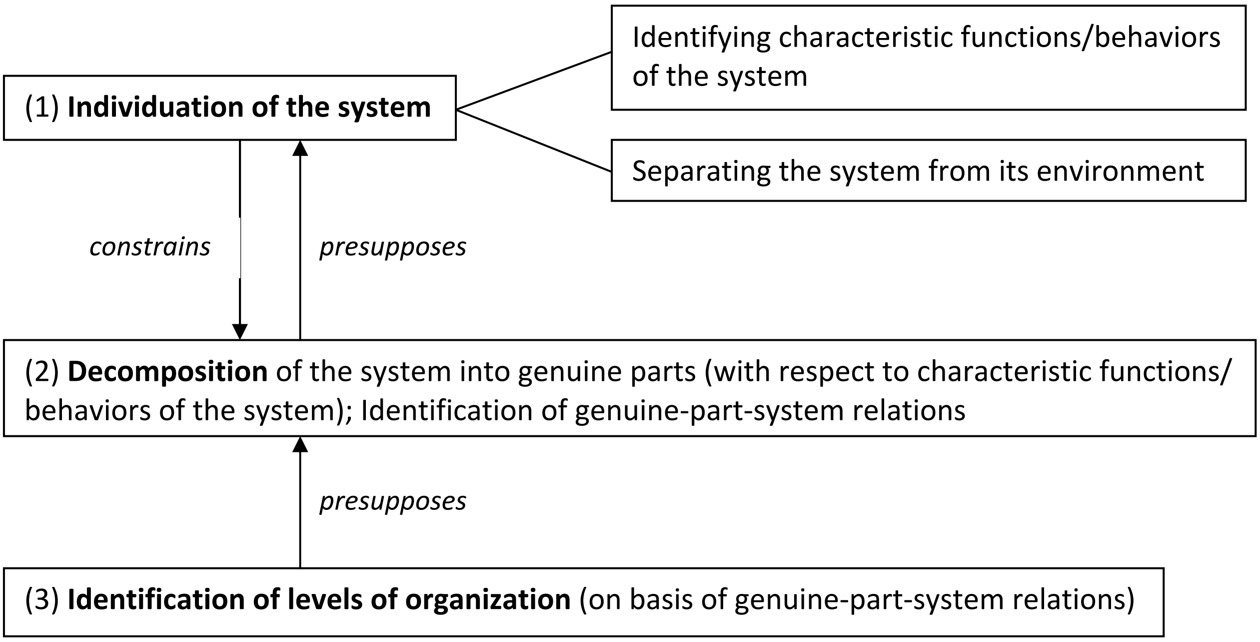

Figure V.11: The Three Tasks of Investigating Biological Systems

In the remaining, second part of this chapter I will apply the concepts and distinctions that I have introduced and developed in this section to examine what it is that makes biological explanations reductive. Roughly speaking, my central thesis will be that reductive explanations exhibit three characteristics: first, they exclusively refer to lower-level factors (Section 2), second, they focus on factors that are internal to the system (Section 3), and third, they appeal only to parts in isolation (Section 4). After disclosing these three features I will conclude by highlighting the differences between reductive explanation and part-whole/ mechanistic explanation (Section 5) and by spelling out the ontic character of my account of explanatory reduction (Section 6).

2 Lower-level Character

In this section I introduce and discuss one of the three characteristics that make biological explanations reductive, namely their lower-level character. The term ‘lower-level character’ refers to the feature of reductive explanations that they explain the behavior of a system by *appealing only to lower-level factors* (more precisely, to factors that are located on at least one level below the level of the system whose behavior is to be explained). The result of my analysis will be that the lower-level character is a *necessary condition* for a biological explanation to be reductive, which, however, is not sufficient on its own.

I start with analyzing paradigmatic and important examples of reductive explanations that stem from that biological discipline that is regarded as the embodiment of the success of reductionism, namely molecular biology (Section 2.1). In Section 2.2 I scrutinize the lower-level character of reductive explanations. I point out that they involve a unidirectional flow of explanation from the bottom or lower level to the top or higher level, which is why the appeal to downward causation renders an explanation non-reductive (Section 2.2.1). Section 2.2.2 serves to specify what exactly it means that explanations refer only to lower-level factors. This involves answering at least two questions: first, does the distinction between lower-level and higher-level factors coincide with the distinction between genuine parts and non-genuine parts of a system? (As you might expect on the grounds of my remarks on the notion of level, I will deny this.) Second, do functional properties of parts and the organization of parts count as lower-level factors? In other words, does reference to functional properties of parts or to the way they are spatially and temporally organized violate the lower-level character of an explanation? Finally, in Section 2.3 I introduce two significant subtypes of lower-level explanations, namely fundamental-level explanations, which I characterize as molecular explanations that encompass genetic explanations (Section 2.3.1), and single-factor explanations (Section 2.3.2). My central thesis in this section will be that philosophers should be cautious not to identify reductive explanations with fundamental-level explanations because this obscures the diversity of reductive explanation in biological science and converts explanatory reductionism into a straw man that is too easy to reject.

2.1 Starting with Molecular Biology

The message of the subtitle of Rosenberg's book "Darwinian Reductionism – Or, How to Stop Worrying and Love Molecular Biology" (2006) is: if you want to become a reductionist with respect to biology start loving molecular biology. For Rosenberg molecular biology is *the fundamental* biological discipline in the sense that it (and it alone) provides the resources for explaining *all* biological phenomena. Most philosophers of biology disagree with Rosenberg's radical version of explanatory reductionism. However, they concur that the impressive growth and development of molecular biology during the last 60 years motivated a great deal of the reductionist thinking in and about biology.[39] Since James D. Watson and Francis H. C. Crick (1953) discovered the double helix structure of DNA molecular biology not only has solved all its classical problems (e.g. the role of genes in heredity and the relationship between genes and their molecular products; cf. Darden/Tabery 2010), but it has also been

---

[39] Also in biology itself molecular biology is frequently characterized as the "triumph of reductionism" (Fang/Casadevall 2011, 1402).

expanded to many other biological fields. Cell biology, developmental biology, genetics, and immunology are just some of the disciplines that "went molecular" (Darden/Tabery 2010). Nowadays molecular techniques are utilized in almost all biological disciplines.

In the 1970s and 80s the focus of research in molecular biology shifted to the genome and its sequence of nucleic acid bases. Various experimental and computational techniques for producing, storing, and interpreting sequence data of important (model) organisms were developed. The success of genomics culminated in the discovery of the sequence of the human genome a decade ago (the "human genome project", see Venter, et al. 2001). The promise is that DNA sequence information can provide foundations for solving many important scientific and medical problems.

Nowadays many biologists speak about the dawn of a new era, i.e. the *post-genomic era*, which often goes along with a criticism of traditional reductionist approaches (some of which will be addressed in Section 4). However, the immense success of molecular biology during the last 60 years is undeniable. In light of this success story it is not surprising that several biologists and philosophers assign a great explanatory scope and reducing power to molecular biology. In their eyes, knowledge about molecules (and genes) is the clue to explaining phenomena from (almost) all biological areas, ranging from protein synthesis, over heart development in vertebrates and sickle cell anemia, to species coexistence and to the evolutionary origins of certain trait. In Chapter III, Section 1.3 and in Chapter IV, Section 2.2 I indicated some deep problems such a radical version of explanatory reductionism encounters. Nevertheless, I will not take a stand here as to whether this confidence in the explanatory and reducing power of molecular biology is justified or not. What is of interest to me is that, given its immense success in various biological fields, molecular biology seems to be a good place to start looking for paradigmatic cases of reductive explanations.

Example: DNA Replication (Molecular Biology)

Let us start with looking at the solution that was offered to one of the classical problems in molecular biology, namely to the problem of understanding the process (or mechanism) by which the genetic information in a cell is accurately duplicated during cell division, known as *DNA replication*. The explanandum phenomenon is a certain behavior of a cell (i.e. the duplication of its genome) and it is explained by reference to certain components of the cell and the various interactions between them. To put it in a nutshell, the DNA double helix, which is a nucleic acid composed of a sugar-phosphate backbone and nucleic acid bases, unwinds (with the help of DNA helicases and topoisomerases) and each of the two DNA strands serves as a template for synthesizing a new DNA strand. DNA polymerases catalyze the synthesis of the new strands by lowering the activation energy for forming covalent bounds between the building blocks of the new strands, the nucleotides. A DNA polymerase can only synthesize a new DNA strand in 5'-to-3' direction and since the replication fork moves in this direction only on the leading strand only the synthesis of this new DNA strand is continuous. The synthesis of the other new DNA strand (with the lagging strand as template) proceeds piecemeal and the emerging Okazaki fragments are connected afterwards. The old and the new DNA strands become connected via hydrogen bonds between the nucleic acid bases (adenine pairs with thymine, guanine with cytosine). In fact, the process of DNA replication is even more complicated and involves the interactions of further entities (e.g. DNA primases, single-strand binding proteins, etc.), as the subsequent figure indicates.

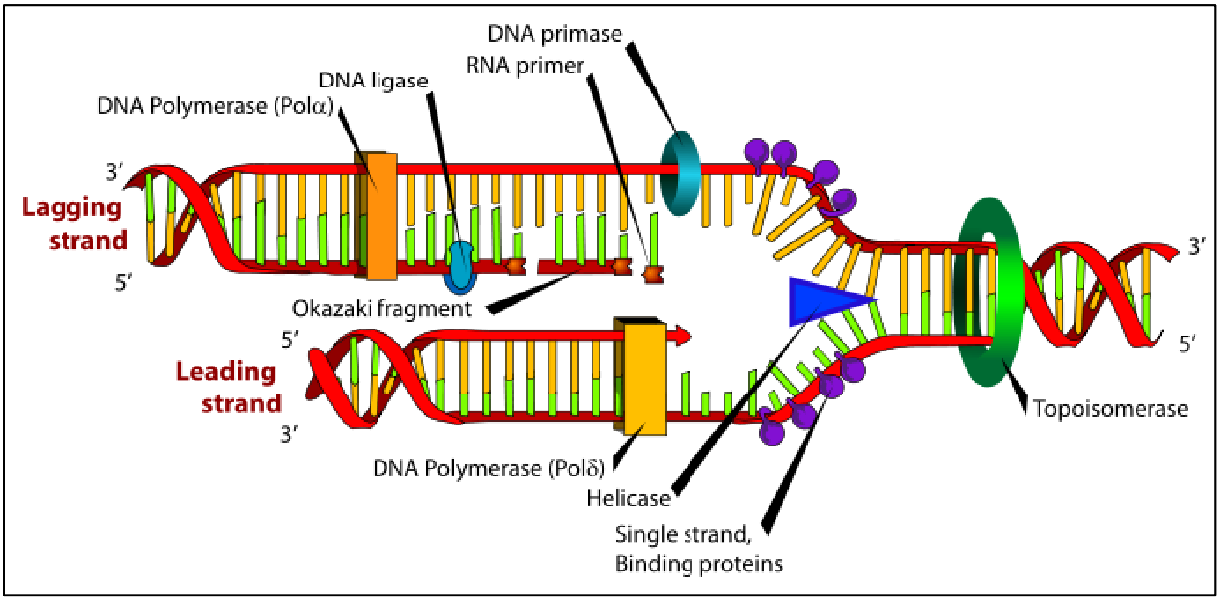

Figure V.12: DNA Replication

Before I analyze the reductive character of this explanation, let us consult a second example. It stems not from the core area of molecular biology itself, but from a field which has been "molecularized" in large parts, namely from plant physiology.

Example: Photosynthesis (Plant Physiology)

Photosynthesis is the process (or mechanism) by which plants (as well as algae and some species of bacteria) convert carbon dioxide into organic compounds, especially sugars, using the energy from sunlight. If one focuses on the first part of this process, known as light reaction, the phenomenon to be explained is the behavior of chloroplasts to transform light energy into chemical energy (in the form of ATP and NADPH). This phenomenon is explained by a certain sequence of sub processes that take place in and next to the thylakoid membrane in the chloroplasts. First, the chlorophyll molecules in photosystem II absorb photons whereby an electron is excited and picked up by an electron transport chain. The remaining electron-deficient hole in the reaction center of photosystem II is filled by an electron originating from the splitting of water. As the electrons are transported down the electron transport chain, (along the way their energy level is raised a second time through the absorption of photons in photosystem I), released energy is used to pump hydrogen ions (H^+) across the thylakoid membrane. This creates an electrochemical proton gradient which is, in the last step of light reaction, used (by an ATP synthase) to generate ATP. The end of the (non-cyclic) electron transport chain is reached when the electron is transferred to $NADP^+$ whereby NADPH is produced.

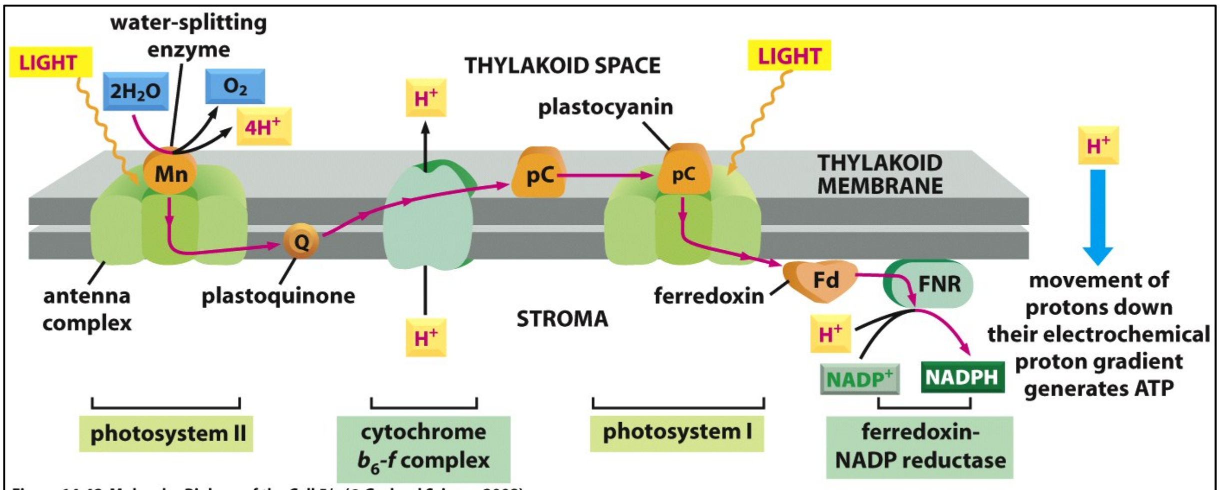

Figure 14-48 Molecular Biology of the Cell 5/e (© Garland Science 2008)

Figure V.13: Light Reaction of Photosynthesis (Alberts, et al. 2008, 852)

What are the features due to which these two explanations are characterized as reductive? One answer is that these explanations both explain a particular phenomenon or behavior of a system by reference to factors that are located on a lower level than the explanandum phenomenon. To specify the lower-level character of reductive explanations is the central goal of this section.

2.2 Lower-level Explanation

The two examples of molecular explanations illustrate a major characteristic of reductive explanations in biology: to explain a phenomenon (or a behavior of a system) reductively means to explain it by reference exclusively to factors that are located on a *lower level* of organization than the phenomenon to be explained. In other words, in the case of reductive explanations there exists a specific relation between the entities referred to in the explanandum and those referred to in the explanans, namely the latter are located on a lower level of nature than the former.[40] For instance, the behavior of a cell to duplicate its genetic information is reductively explained by describing the organization of and interactions between certain components of the cell, such as DNA strands being separated by helicases, DNA polymerases synthesizing new DNA strands, etc. Likewise, the behavior of chloroplasts to transform light energy into chemical energy is reductively explained by reference to lower-level parts, such as photosystems, electrons, electrochemical proton gradients, and ATP synthases. Thus, reductive explanations *bridge* at least *two levels*, namely

---

[40] This is also called the *interlevel character* of reductive/mechanistic explanations (e.g. Brigandt/Love 2008, Section 5; Bechtel 2008, 148).

the higher level of the explanandum phenomenon and the lower level(s), on which the explanatory relevant factors cited in the explanans are located on.[41] The term ‘at least’ indicates the possibility that, in addition to the explanatory relevant factors being located on a lower level than the explanandum phenomenon, the explanans itself can appeal to entities from different lower levels. For instance, one may argue that a photosystem (i.e. a complex of macromolecules) is not located on a “higher” lower level than an electron (i.e. a subatomic particle) because photosystems are, among other things, composed of electrons. However, unlike the lower-level character, the explanans of reductive explanations must not involve *different* lower levels.[42]

2.2.1 Unidirectional Flow of Explanation

Since the direction of reductive explanations runs from the lower level to the higher level biologists also refer to them as “*bottom-up [explanations]*” (Sorger 2005, 9; Soto/ Sonnenschein 2009, 5; my emphasis) or “upward explanation[s]” (Mikkelson 2004, 120; Lidicker 1988, 278).[43] This unidirectional flow of explanation from the bottom (i.e. lower level) to the top (i.e. higher level) is also called “’level’ directionality” (Schaffner 1993, 431). For instance, in the reductive explanation of DNA replication the explanation proceeds from the lower level to higher level. That is, the description of a specific sequence of particular interactions between macromolecules, that is, DNA strands and various proteins (lower-level), explains how a cell duplicates its genome during meiosis (higher-level), not the other way round.

The characterization of reductive explanations as bottom-up or upward explanations makes it intelligible why the existence of *downward causation* would render the explanation of a phenomenon produced by downward causation non-reductive. This is not to say that downward causation definitely exists. Rather, my point is that if there were phenomena that were (at least partly) causally affected by higher-level entities then the causal explanation of this phenomenon would not be reductive because the flow of explanation would (at least partly) run from the top (i.e. higher levels) to the bottom (i.e. lower levels).[44] In cancer

---

[41] This thesis should not be confused with the view of reduction that is known as “*layer-cake reduction*” (e.g. Rosenberg 2006, 28; Waters 2008; Steel 2004, 60) and which can be traced back to Oppenheim’s and Putnam’s layer-cake model of the unity of science (1948). My account of reduction is neither concerned with theory reduction (see Chapter II, Section 3.1), nor does it accept Oppenheim’s and Putnam’s concept of level (see Section 1.3.1).

[42] The latter is commonly referred to as the “multilevel character” of explanations (e.g. Mitchell 2009, 109-115; Schaffner 2006, 384-387; Craver 2007a, 9-16; see also Section 1.3.3).

[43] Note that the direction of a reductive explanation runs in the opposite direction than the reductive method decomposition. Systems are decomposed into their parts by looking down one or more levels, whereas the behavior of a system is explained by looking upward.

[44] Mikkelson uses this characterization of reductive explanation to offer a quite moderate view of the reductionist position: “The reductionist considers ‘upward’ causal and explanatory relationships to be more important than ‘downward’ causation and explanation.” According to his view, only a “strict reductionist… believes that downward explanations are illegitimate, and should therefore be expunged altogether from science.” (2004, 120f)

research this difference between, as they call it, the "reductionist" and the "organicist/ systemic approach" (Soto/Sonnenschein 2009, 6) is illustrated by the following figure:

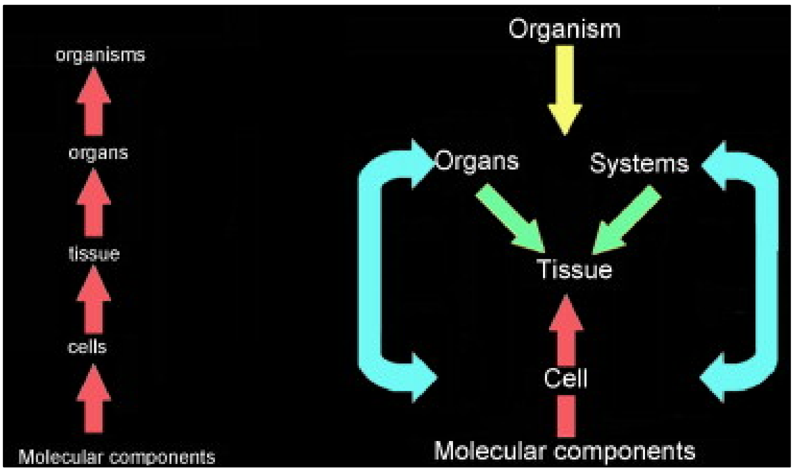

Figure V.14: Reductionist vs. Systemic Approaches in Cancer Research (Soto/Sonnenschein 2009, 6)

According to Soto and Sonnenschein, reductionists assume that causal relations between entities from different levels exclusively run from lower to higher levels, whereas proponents of an organicist/systemic approach allow causal relations also to run from higher to lower levels (or in both directions).[45] As neat as this picture might be, it must be treated with care since the assumption of the existence of *interlevel causal relations*, in particular of *intrasystemic* interlevel causal relations (i.e. causal relations between wholes and their parts), is problematic (see e.g. Craver/Bechtel 2007). Thus, it is questionable whether the arrows in the above figure can be interpreted as causal relations. Instead, one could argue that they represent relations of composition (the red arrows) or of (co-)determination (the yellow and green ones). For instance, typically we would say that cells compose tissues, but not that they cause them. Likewise, the behavior of an organism as a whole (e.g. that it is running fast) may determine the behavior of the organs of which it is composed (e.g. that the heart beats fast). But do we really want to say that the organism causally affects its parts? I doubt this. However, this is not a definite judgment since the thesis of interlevel causation needs far more treatment that I can provide it here. Whether or not interlevel causation exists does not affect my thesis that reductive explanations are to be characterized as bottom-up or upward explanations since the description of bottom-up causation is not the only thing that gives an explanation a clear direction from lower to higher levels (for instance, the description of constitutive relations suffices, too).

Despite the problematic reliance on downward causation, Soto's and Sonnenschein's argument has a virtue, too. It discloses a *possible line of criticism* against reductive explanation. A common antireductionist objection is to argue that adequate explanations of

---

[45] In a different paper Soto and Sonnenschein explain this difference as follows: "In this view, causes act from the bottom-up. Contrary to reductionism, organicism considers both bottom-up and top-down causation." (2010, 364)

(many or certain) biological phenomena do not exhibit a unidirectional flow of explanation. That is, antireductionists frequently claim that adequate explanations of biological phenomena do not merely flow from lower to higher levels, but also the other way round. In line with this, for instance, Kitcher states that the understanding of a phenotypic trait requires a "constant shifting back and forth across levels" (1984, 371).

2.2.2 Exclusion of Higher-level Factors

Another way to frame the lower-level character of reductive explanations is to say that the explanantia of reductive explanations do not appeal to higher-level factors. This gives rise to the question of what exactly it means that higher-level factors are excluded from the explanans. In other words, how can the distinction between lower-level and higher-level factors be specified? In what follows I discuss two suggestions how to draw this distinction.

Why Lower-level Factors Are More than Genuine Parts

First, one might be faintly reminiscent of my analysis of the concept of level in Section 1.3 and propose to identify a higher-level factor with the system whose behavior is to be explained and lower-level factors with the parts of this system. According to this reading, a reductive explanation of the behavior of a system would appeal exclusively to the *parts* of this system. Thus, the class of reductive explanations would be identical to the class of *part-whole explanations*. This proposal has been made for instance by Hüttemann and Love (2011; see Chapter III, Section 2.3). There is much that is right about this suggestion. As we have seen in Section 1.3, the notion of levels of organization is closely related to the notion of part-whole relations. More precisely, I claimed that a certain kind of compositional relations, namely genuine-part-system relations, determine which levels are to be distinguished in the biological realm. However, it is also important to note in which respects it is misleading simply to equate lower-level factors with genuine parts (of a particular system). The core of the problem is that although the genuine parts of a system definitely are located on a lower level than the system, the genuine parts are not the only factors that are "lower-level" (relative to the system of interest). However, this argument requires the *notion of a global level*, rather than just a concept of level, according to which levels can be defined only locally. To explain why, I must go into greater detail.

According to the concept of a locally defined level, such as Craver's notion of mechanistic levels, claims about levels can only be made with respect to the components of a mechanism (or with respect to the parts of a system). The question of whether factors that are external to the mechanism/system in question are located on a higher level, on a lower level, or on the same level than the mechanism/system simply cannot be answered (or "makes no sense"; Craver 2007a, 191). In other words, according to a local notion of level, claims about levels are restricted to mechanisms and their components (or, generally speaking, to systems and their parts). Contrary to this, in the biological literature on reductionism and on reductive explanation we frequently find assertions such as that citing contextual factors does not render an explanation non-reductive, provided that the context is described in lower-level terms. In philosophy, too, the strategy to "reduce the context", that is, to spell out the context in lower-level terms, is seen as a permissible strategy in order to defend reductionism (e.g. Delehanty 2005).[46] These claims presuppose that we can assess whether contextual factors are located on a higher level, on a lower level, or on the same level as the system of concern. In other words, these claims demand a concept of global levels. In Section 1.3 I suggested that a global notion of level might be obtained by bringing together different genuine-part-system distinctions and establishing a hierarchy of different genuine-part-system relations.

On the basis of such an account of global levels we also get an idea of why reductive explanations are not identical to part-whole explanations (this difference will be further clarified in Section 5).[47] In my view, the class of reductive explanations is *broader* than the class of part-whole explanations because there exist explanations that are reductive, but that fail to be part-whole explanations. Explanations of this kind appeal to factors, which are external to the explanandum system (i.e. that are no parts of the system), but which are nevertheless located on a lower level than the system. An example is the explanation of how the PI 3-kinase/Akt pathway promotes cell survival (this example is discussed in detail in Section 3.3). This explanation describes how several proteins step by step are phosphorylated and thereby activated, and how this finally causes the release of inhibitory proteins, which block apoptosis. Most of the interacting entities the explanation refers to are genuine parts of the cell, but not all of them. For instance, the explanation starts with the description of the binding of an extracellular survival signal to a particular receptor of the cell. This factor is external to the cell, but at the same time it is a lower-level factor since it belongs to the class of macromolecules, which frequently are genuine parts of cells.

To conclude, the concept of a local level is insufficient because it does not allow making claims about the levels of factors that are external to a system of interest. If we adopt an account of global levels we realize that not only are all genuine parts of a system are located on a lower level than the system, but that also contextual factors can be said to be lower-level factors. Consequently, the distinction between higher-level and lower-level factors only partially coincides with the distinction between parts and non-parts.

Do Organization and Functional Properties Count as "Lower-level"?

Second, let us assume for the sake of argument that the term 'being located on a lower level than system A' has the same extension as the term 'being a genuine part of system A'. Even if we accepted this, the question of what these lower-level factors are that are exclusively

---

[46] This is not to say that this strategy can always be successfully conducted.

[47] This is a crucial difference between my account of explanatory reduction and Hüttemann's and Love's analysis (2011). They identify reductive explanation with part-whole explanation; whereas I distinguish these two kinds of explanation (see also Chapter III, Section 2.3).

referred to in reductive explanations would still not be fully answered. For instance, it would remain an open question whether the *organization* of the genuine parts of a system and the *interactions* between them count as lower-level factors, too. Furthermore, it can be disputed whether all kinds of properties of genuine parts belong to the lower-level factors that are allowed in reductive explanations, for example, whether *functional properties* of parts are allowed or just structural ones.[48] These questions are closely linked to issues that take center stage in Section 4. This is why they will not be fully examined in this section, but will be taken up again in Section 4. Let us start with analyzing an example from developmental biology as well as the philosophical discussions that have been launched about it.

Example: Limb Development of Vertebrates (Developmental Biology)

The formation of the vertebrate limb is an instance of *pattern formation* (i.e. of the development of an animal's specific arrangement of organs and tissues in their characteristic places in three-dimensional space). The development of a limb starts with the formation of a *limb bud*. Two regions in a limb bud have profound effects on the limb's development, namely the *apical ectodermal ridge* (AER) and the *zone of polarizing activity* (ZPA).

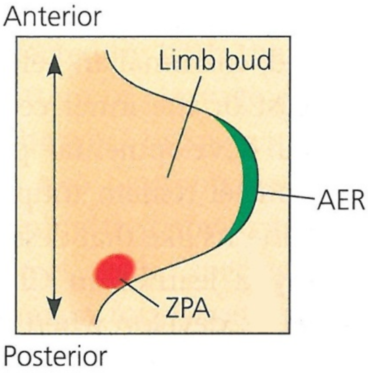

Figure V.15: Formation of a Limb Bud (Reece, et al. 2011, 1086)

The cells of these regions secrete proteins that provide key positional information to the other cells of the bud. The positional information indicates location along three axes: the proximal-distal axis, the anterior-posterior axis, and the dorsal-ventral axis. For instance, ZPA is necessary for proper pattern formation along the anterior-posterior axis (i.e. the "thumb-to-little finger" axis) of the limb. Cells nearest the ZPA give rise to posterior structures; cells farthest from the ZPA form anterior structures.[49]

---

[48] Already Wimsatt has recognized the importance of questions like these. He states that "the issue between scientists who are reductionists and holists is not over the in principle possibility of an analysis in lower-level terms but on the complexity and scope of the properties and analyses required. The more holistically inclined scientists usually argue that higher-order relational properties of the lower-level entities are required, and the reductionists argue that a given simple, lower-level model... is adequate." (1980, 146)

[49] For detailed information about limb development see Gilbert 2006, Chapter 16.

This example has been picked up on by Kitcher and invoked as evidence against reductionism (1984, 371-373). More precisely, Kitcher is concerned with a mutation that leads to the disruption of the normal gene expression in limb development. That is, the explanandum phenomenon he is interested in is the non-expression of certain genes that are important for the development of a functioning limb. This phenomenon is explained by describing how a mutation of a limb-bud allele gives rise to non-functional proteins, which cause peculiarities of cell shape. This, in turn, affects the spatial relations among the cells of the embryo (i.e. it causes an abnormal tissue geometry). Due to the gaps between certain cells the protein secreting regions (i.e. AER or ZPA) are too far away from the target cells. That is, in those cells whose genes normally should be expressed the concentration of the protein gradient is below the level required for gene expression, which is why these genes are not being expressed.

From Kitcher's point of view, this explanation of the non-expression of specific genes in limb development supports antireductionism because it involves describing how the presence of certain higher-level factors (i.e. the gaps between cells that are normally adjacent) gives rise to changes in gene expression. In his words,

> claims at a more fundamental level (specifically, claims about gene expression) are to be explained in terms of claims at a less fundamental level (specifically, descriptions of the relative positions of pertinent cells). (1984, 372)

What Kitcher seems to state is that the explanation at hand is non-reductive because it appeals not only to lower-level factors (such as mutant alleles and non-functional proteins), but also refers to higher-level factors (such as the relative position of cells or tissue geometry).

Rosenberg denies this. He argues that in the explanation of abnormal gene expression in limb development higher-level terms (such as 'cell') are dispensable and can easily be replaced by macromolecular terms (such as 'lipid bilayer').[50] In particular, Rosenberg states that "the effect that the presence of the cell has on development is identical to the effect that a purely macromolecular structure has" (2006, 83). He concludes that we can explain the non-expression of genes without invoking the higher-level term 'cell'. Instead, it suffices to refer to the gap between lipid layers and to explain that due to this gap the distance from AER or ZPA to particular genes is so far that the concentration of the protein gradient is below the level required for gene expression. An antireductionist like

---

[50] Since this sounds very eliminative it should be stressed that Rosenberg seeks to defend a non-eliminative version of explanatory reductionism in biology. However, in Chapter III, Section 1 I have already pointed out that it remains unclear what exactly this non-eliminative character amounts to. On the one hand, Rosenberg claims that reductionism tolerates "terms like *cell* as acceptable expressions in biological description" and "accepts the reality of cells and their causal roles". On the other hand, he denies that "there are distinct causal properties of the items such terms name that are not open to identification in molecular terms" (2006, 84). Other statements, too, suggest that Rosenberg assigns only a very minor role to functional biology's concepts (such as 'cell'). In his view, they may play an important role for identifying the phenomena to be explained, but they are not part of the explanantia (see Chapter III, Section 1.1.2).

Kitcher might counter that an explanation which dispenses all higher-level terms is inadequate. He might insist that higher-level terms like 'cell' are not dispensable, but rather are essential parts of adequate explanations. What interests me about this discussion is not whether the provided arguments in favor of or against explanatory reductionism are more conclusive (some of these issues have been addressed in Chapter IV, Section 2.2). Rather, my central concern is how these authors spell out which kinds of objects and properties are lower-level factors and, thus, allowed in reductive explanations.

The just presented discussion about how to explain abnormal gene expression in limb development provides some insights. The phenomenon to be explained is gene expression, which might be characterized as a process in which genes and other molecules are involved, or as a property (or event) of genes. In any case the phenomenon to be explained is located on the *molecular level*. Accordingly, if the explanation refers to cells it appeals to objects that are located on a *higher level* than molecules. The same holds for the gaps between cells, the relative position of cells, and the distance between cells and certain protein secreting regions (e.g. AER or ZPA), which all might be characterized as relational properties of cells. Abnormal tissue geometry, which is also mentioned in the explanation (at least in Kitcher's), is a property of a whole compound of cells, namely of a tissue. That is, it is located on a level that is still higher than the level of cells (thus, it is even more a higher-level factor). However, Rosenberg's and Kitcher's discussion does not help to answer the question of whether the organization and functional properties of parts count as lower-level, too. In order to tackle this question, consider the contributions that other authors have made to this discussion.

In his earlier paper on "Reductionism Redux: Computing the Embryo" (1997) Rosenberg presents a less well-developed version of his argument than I outlined above. This paper initiated a debate about reductive explanation in developmental biology. The involved authors (most notably, Laubichler/Wagner 2001 and Frost-Arnold 2004) address questions such as: can the *organization* of molecules be described in molecular terms (i.e. in the reducing language)? Is it possible to assign *functions* to genes and molecules in a reductive explanation? Which spatial terms are allowed in the *reducing vocabulary*? That is, they explicitly approach the issue of how to distinguish what counts as lower-level and thus belongs to the reducing realm.

Let us start with the question of whether reference to *functional properties* of molecules or genes violates the reductive character of explanations. Note that the discussion of Rosenberg, Laubichler, Wagner, and Frost-Arnold focuses on reductive explanations that are molecular explanations, that is, in which the lower level (or reducing realm) is the level of molecules.[51] Accordingly, the question is whether we can refer to functional properties of

---

[51] One should be aware of the fact that not all reductive explanations in developmental biology must bottom out at the molecular level. For instance, gastrulation can be described as a process in which the whole blastula is involved. Gastrulation of the blastula is explained by appealing to certain regions of the blastula and to the tree germ layers. These explanatory relevant factors are genuine parts of the blastula and located on a lower level than it, but they are not molecules.

molecules and genes without turning the explanation into a non-molecular (and thus non-reductive) explanation. Laubichler and Wagner deny this. They argue that we can only assign well-defined functions to genes in a developmental process if we take into account "the larger molecular, cellular, and organismal context within which these genes are expressed" (2001, 59). The reason they adduce is that the function of a gene does not only depend on its nucleotide sequence, but on various other factors that are present in its context. Likewise, the specificity of a protein is not only determined by its amino acid sequence, but also by different factors that are part of its environment.[52] In short, there is no "simplistic, linear, causal relationship 'gene – protein – function'" (Bizzarri, et al. 2008, 181). Frequently the same molecule has different effects when it is located in different contexts. For instance, the sonic hedgehog protein, which plays a key role in regulating vertebrate organogenesis, such as in the growth of digits on limbs and organization of the brain, activates different proteins in different contexts or morphogenetic fields (Laubichler/Wagner 2001, 63). But what follows from this for the reductivity of an explanation? Does an explanation that appeals to functional properties of genes and molecules still exhibit a lower-level character?

In my view, reference to functional properties of parts does not violate the lower-level character of an explanation, unless the "contextual factors" that need to be included in the explanation cannot be spelled out in lower-level (i.e. in molecular) terms, too (see also Frost-Arnold 2004, 81-84). In other words, molecular explanations can refer to the functions of genes and molecules without becoming non-molecular/non-reductive if the "contextual factors" on which these functions depend (and which need to be described in the explanation) are spelled out in molecular/lower-level terms as well.[53] However, two things should be noted at this point: first, the notion of contextual factors involved here is different from the one I will introduce in Section 3 when I analyze the internal character of reductive explanations. Here the context *of the system's parts* is meant, which may include only intrasystemic factors (i.e. factors that are located inside of the system in question), or which may (but need not) include parts of the context *of the system*, too. If the functional properties of the system's parts are specified by appealing to lower-level factors that are external to the system, the lower-level character of this explanation will not violated, but it's internal character will be (for further details see Section 3). Second, it is important not to confuse two questions: on the one hand, the question that is at issue here, namely the question of whether the inclusion of functional properties of the parts of a system violates

---

[52] Van Regenmortel explains this as follows: "the specificity of a complex biological activity does not arise from the specificity of the individual molecules that are involved, as these components frequently function in many different processes. […] It is the particular cellular compartment and environment… that allow a gene product to have a unique effect. Biological specificity results from the way in which these components assemble and function together." (2004b, 1016; see also van Regenmortel 1998)

[53] This is compatible with the view that there might exist an *additional feature* of reductive explanation, according to which at least in some reductive explanations structural properties are regarded as being more important than functional ones. This view can be found, for instance, in Wimsatt's work (e.g. 2006a, 468) and it is implicitly present in Sarkar's analysis of reductive explanation in genetics and molecular biology (e.g. 2005, 9f). However, I do not endorse this view because it is not supported by the biological literature.

the reductive character of the explanation of the system's behavior; on the other hand, the question of whether functional properties *of biological systems itself* can be explained in a reductive manner or not.[54] An affirmative answer to the former does not imply any kind of answer to the latter.

Finally, consider what Laubichler and Wagner say about the organization of molecules. They accuse Rosenberg of overlooking that "the vastly complex combinations of molecules that make up biological objects are *organized* and *structured*" (2001, 58). Furthermore, they claim that "the *organization* of these molecular building blocks into morphological structures is not part of the molecular syntax" (2001, 62). I take them to be saying that the description of the temporal and spatial organization of parts into a higher-level system cannot be part of the explanans of a reductive explanation. The corresponding objection that reductive explanations are inadequate because they do not account for the organization (or "integration"[55]) of parts into a system is a common line of criticism of antireductionists. However, the question remains whether it is cogent to treat reductive explanation as such explanations that neglect the organization of lower-level factors all together. This question will be approached in detail in Section 4 when I examine an additional feature of reductive explanation. What matters at this point is whether the *lower-level character* of reductive explanations, as such, precludes that the organization of lower-level factors is described in reductive explanations. Laubichler and Wagner offer an affirmative answer. They claim that organization is not part of the "molecular syntax" (i.e. the reducing vocabulary). By contrast, I think we need to endorse a more balanced view.

First of all, what decides whether organizational features can be represented in a reductive explanation is whether these organizational features are to be characterized as properties *of lower-level objects* (probably as relational properties of the system's parts) or as properties *of the system* whose behavior is to be explained (i.e. as systemic properties). Lower-level properties can be referred to in a reductive explanation, whereas reference to systemic properties violates the reductive character of an explanation. Consider an example. The explanation of anterior-posterior axis formation in early chick development (e.g. Gilbert 2006, 343-347) appeals to forces of gravity and adhesion that act on the blastodisc. These forces establish a top and bottom of the blastodisc, which causes axis determination since the molecules located at the top are different from those between molecules located at the bottom. It seems to me that this explanation involves spatial predicates of two kinds (see also Frost-Arnold 2004, 85-89): on the one hand, there are spatial predicates at the molecular (i.e. lower) level (such as 'molecule A is located close to molecule B'). On the other hand, the explanation appears to involve also embryo-level spatial predicates (such as 'being

---

[54] With regard to the latter question see, for instance, Sarkar 2005, 117-143.

[55] "[G]ene sequencing and other techniques will soon have isolated all the cell's individual parts and spelled out their isolated functions. Now, it is time to move beyond reductionism. […] Now we need to know how all these things are *integrated*." (Service 1999, 81)

located at the bottom of the embryo'). Only the latter of them are so called *systemic properties* whose inclusion renders an explanation non-reductive.

To summarize, the purpose of this section was to state more precisely what lower-level factors are and whether functional properties of parts and the organization of parts belong to the class of lower-level factors, too. The result of my analysis is threefold: first, the set of factors that are located on a lower level than a particular system of interest is not restricted to the genuine parts of this system. Rather, factors that are external to the system, but are located on the same level(s) than the genuine parts of the system, are lower-level, too. Second, the class of lower-level factors also encompasses functional properties of lower-level objects, as long as contextual factors that need to be included as well are spelled out in lower-level terms. Third, if the organization of the parts of the system can be conceptualized as lower-level (relational) properties and not as systemic properties, it will count as lower-level factor, too.

2.3 Subtypes of Lower-level Explanation

The examples analyzed so far all stem from molecular biology or from fields that have been "molecularized" to a great extent. As a result, these explanations all "bottom out" (Machamer, et al. 2000, 13) at the molecular level. In the following section I identify molecular explanations as an important subtype of reductive explanations. However, it must be emphasized that reductive explanations *need not* go down to the lowest biological level of molecules[56] (which I call the *fundamental* biological level). Rather, reductive explanatorily practice in the biological sciences shows that there exist a *myriad of reducing levels*. What matters for the reductive character of an explanation is that the explanatory relevant factors are located on a level that is at least *one level below* the level of the explanandum phenomenon. But this lower level need not be the molecular level. In order to illustrate this, consider an example of a reductive explanation from ecology.

Example: Food Web Structure Explained by IBMs (Community Ecology)

A recent development in ecology is to use so called individual-based models (IBMs) to explain community-level properties, such as the structure of food webs. IBMs represent a community (or a population) as a collection of individual organisms with variable properties. Moreover, the interactions between individuals of different species are incorporated in

---

[56] I am aware that this way of talking about the *one* molecular level is exposed to criticism. The already discussed example of photosynthesis shows that the notion of molecules encompasses objects that are of very different kind and size (e.g. complexes of macromolecules, ions, and subatomic particles) and that partly compose each other (e.g. electrons are parts of photosystems). However, the fact that the molecular level (in a wide sense) can be characterized in a more fine-grained manner by dividing it into the level of complexes of macromolecules, the level of macromolecules, the level of molecules (in a narrow sense), the level of atoms and so on does not imply that the way of talking about the one molecular level (which can be found frequently in biological practice) is illegitimate.

these models. IBMs are typically studied by simulation on a computer. Have a look at a more specific example. Spencer (1997) constructs an IBM for small, freshwater, benthic habitats with three trophic levels (algae, herbivorous invertebrates, and predatory invertebrates). He concludes from his simulations that patterns of food web structure in habitats of different sizes are likely to be the result of a combination of island biographic effects and spatial effects on predator-prey systems (1997, 311). IBMs like the one Spencer construes do not only give rise to predictions, but also to explanations. These explanations are reductive because they explain the properties or behavior of a community by reference merely to individual organisms and the interactions between them (Sarkar 2009, Section 5).[57] For instance, Spencer explains the particular dynamics of the food web structure of the studied community (e.g. the increase of the number of species, the persistence of the predator-prey relations, changes in the number of autotrophs, etc.) by certain interactions between individual organisms from different groups (e.g. that certain individuals diminish the productivity of others, that they immigrate, that they eat others, etc.).

The ecological explanation clearly shows that the level of molecules is not the only lower level to which reductive explanations in biology appeal. The reducing level is here the level of interacting organisms.[58] Brandon has called the reductionist position which recognizes the variety of different reducing levels in the biological science "*multi-level reductionism*" (1996, 182). He contrasts it with "single-level reductionism" (1996, 187), which holds that there exists a single fundamental explanatory level on which every biological phenomenon can be explained. As I have argued in Chapter IV, Section 2.2, single-level reductionism (to which, for example, Rosenberg's Darwinian reductionism belongs) relies on an unconvincing view of explanation. Here I want to add that single-level reductionism gives rise to a *one-dimensional view of reductive explanation*, too. Reductive explanation is more than just explaining a phenomenon in molecular terms. On which lower level a reductive explanation bottoms out, depends on how the phenomenon to be explained is specified and on other pragmatic factors (for more details see Chapter IV, Section 2.2). To put it into the words of some biologists: "in practice, the reductive thrust stops at the level where 'it makes sense'" (Soto/Sonnenschein 2006, 362).[59]

---

[57] This notion of a reductive explanation in ecology is underpinned by the following statements of ecologists: "reductionism in ecology… see[s] 'true causes' arising at only one level […]. … for reductionists, organisms are the only real objects, while the higher levels of organization are the surface of the truly important events…" (Korfiatis/Stamou 1999, 388) And: "Reductionism… sees the individual species, or ultimately the individuals…, as the only 'real' objects while higher levels are again descriptions of convenience without causal reality." (Levins/Lewontin 1980, 51)

[58] The version of reductionism that corresponds to reductive explanations of this kind is also known as "*methodological individualism*" (Sarkar 2009, Section 5; Dupré 1993, 107-120). In the social sciences methodological individualism prescribes explaining social phenomena by appealing to the behavior of individuals, ignoring or simplifying influences from the environment and from higher levels as the societal level.

[59] The New Mechanists formulates this point as follows: "[T]he components that are accepted as relatively fundamental or taken to be unproblematic for the purposes of a give scientist, research group, or field… [i.e.

2.3.1 Fundamental-level Explanation

As I have already indicated, there exists a common understanding of the concept of a reductive explanation, according to which to explain a biological phenomenon reductively means to explain it in *molecular* terms (sometimes the term 'molecular' is narrowed down to the term 'physicochemical' or 'biochemical'[60]). In other words, reductive explanations are frequently identified with molecular explanations. The following quotations from the biological literature support this claim:

> [B]iological research became largely reductionist (i.e. increasingly involved in the analysis of *molecular* details). (Grizzi/Chiriva-Internati 2005, 29; my emphasis)

> [M]ethodological reductionism… predicates the study of biological systems at the *lowest possible level* with the objective of uncovering *molecular* and *biochemical* causes. (Soto/Sonnenschein 2010, 364; my emphasis)

This understanding is especially common when biologists criticize reductionist approaches:

> The long trend towards analysis of lower and lower levels is starting to reverse. […] [I]t has become obvious that reductionism has its limits. Eventually, the mass of detail from yet *lower levels* of analysis provides no more explanatory power. (Bateson 2005, 31; my emphasis)

Since the level of molecules is regarded as being the *lowest* level of biological organization (of course not of organization in general) I call it the *fundamental level*. This captures an idea that is closely linked to the notion of reduction, namely that there exists a level of basic or ultimate (biological) constituents (Mayr 1988, 475). Reductive explanations that appeal only to factors that are located on this single fundamental level of molecules (i.e. molecular explanations) are also called *fundamental-level explanations*.[61]

Note that the meaning of the term 'fundamental' as I use it here is different from the one that Sarkar, Hüttemann, Love, and the New Mechanists employ (see Chapter III, Section 2.2.3 and 2.3.2). They use the term 'fundamental' in a broader sense, according to which any level on which an explanation bottoms out is "relatively fundamental" (Machamer, et al. 2000, 13). As a result, almost each level of organization counts as (relatively or locally) fundamental, i.e. as fundamental with respect to a phenomenon or a certain field. By contrast, my usage of the concept is more restrictive. I call only the lowest level of biological organization, namely the level of molecules, fundamental. My understanding of 'fundamental' has the advantage that it captures the intuition that a fundament is something of which there is only one and not many.

---

the level, on which] the explanation comes to an end… is relative. […] [And] what is considered as the bottom out level may change" (Machamer, et al. 2000, 13f).

[60] Only a few authors, mostly philosophers of science, speak about biological phenomena being explained in physical terms (e.g. Fodor 1974, 107; Nagel 1952, 17).

[61] Wimsatt describes the reductive method that corresponds to this kind of reductive explanation as follows: "Assume that all descriptions and processes are to be referred to entities at a given level" (2006a, 468).

Genetic Explanations

When we are concerned with the explanation of phenotypic traits of organisms (ranging from the white coat color of mice over the mating behavior of birds to complex diseases in humans, such as depression or cancer) not only molecules, but also *genes* occupy center stage. The central question is whether a particular phenotypic trait of an organism can be explained by appealing merely to the genes of that organism, respectively to "genes and/or [molecular] gene products" (Soto/Sonnenschein 2005, 104; 2006, 362). In past biomedical research much effort has been put in the search for the genetic causes of phenotypic traits. In the 1970s and 80s genomics emerged and the accumulation of sequence data began. Nowadays the DNA sequence of the human (Venter, et al. 2001) and of most important model organisms is known. However, the way from sequence data to the identification of all genes in an organism and to the knowledge of their functions is long. Moreover, in recent years critical voices have increased that accuse biological and medical research of being too reductionistic, that is, of being focused too much on the search for the genetic causes or phenotypic traits. The following quotes illustrate this:

> Five years ago, the number of molecular biologists willing to give up their paradigm of genetic reductionism was still relatively small, but biologists seem to be undergoing a paradigmatic shift right under our noses. […] For 50 years, biologists have focused on reducing life to its constituent parts, first focusing on the cell, then working their way down to the genome itself. However, […] [they] took Humpty Dumpty apart but left the challenge of putting him back together again. (Keller 2005, 4)

> During the last fifty years the dominant stance in experimental biology has been reductionism. […] [G]enes were in 'the driver's seat' controlling the developmental program and determining normalcy and disease (genetic reductionism and genetic determinism). (Soto/Sonnenschein 2005, 103)

These citations exemplify a stance that is prevalent in contemporary biology: previous research is being accused of *genetic reductionism*. By genetic reductionism most biologists mean the thesis that phenotypic traits of organisms most fruitfully are studied at the genetic level, that is, by uncovering their genetic causes (methodological genetic reductionism), and that they are adequately explained by referring to these genetic causes (explanatory genetic reductionism).[62] The shortcomings of genetic reductionism are most extensively discussed in the biomedical sciences itself, first of all in the research field that studies the causes for carcinogenesis (i.e. the genesis of cancer).

Example: Carcinogenesis (Biomedical Science)

Cancer research is one of the most important and heavily funded research areas in the biomedical sciences. Moreover, during the last decade it has generated one of the most

---

[62] The controversial assumption, that the appearance of a particular phenotypic trait of an organism is exclusively determined by its genes, is called *genetic determinism*. This thesis is nowadays rejected by almost everyone (Byerly 2003).

interesting disputes about *conflicting epistemological stances* in biomedical research practice. The most popular discussion turns on the adequacy of reductionism, more precisely on the appropriateness of reductive explanatory strategies.

Soto and Sonnenschein (2000, 2005, 2006, 2009, and 2010) are among the central figures in this dispute. They claim that reductionism has been the prevailing epistemological stance in cancer research, but that it has proved to be deficient and, thus, should be replaced by their "organicist view" (2005, 104). The reductionist approach they criticize is known as the *somatic mutation theory* (SMT). SMT seeks to explain carcinogenesis by appealing to genetic mutations and to the molecular changes in a cell that are affected by these mutations (e.g. Weinberg 1998, Hahn/Weinberg 2002). The organicist view that Soto and Sonnenschein advocate is called *tissue organization field theory* (TOFT). According to TOFT, cancer is a problem of tissue organization (Soto/Sonnenschein 2005, 103), rather than a cellular problem caused by mutated genes.[63]

Let us go into some details. The phenomenon to be explained is the genesis of cancer. Since it is a phenotypic trait that is diagnosed by biopsy it is best characterized as a phenomenon that is located at the tissue level (Soto/Sonnenschein 2005, 112). SMT seeks to *reductively* explain carcinogenesis by identifying those genes in a cell whose mutations cause the tumor progression. The cancer-causing genes are known as "oncogenes" (Reece, et al. 2011, 419). The story that is being told in introductory biology textbooks is fully in line with SMT. Cancer cells divide extensively and invade other tissues. Their abnormal behavior results from certain genetic changes that occur in so called *proto-oncogenes*, which code for proteins that stimulate normal cell growth and division. Different kinds of genetic changes that convert proto-oncogenes to oncogenes can be distinguished: movement of DNA within the genome, amplification of proto-oncogene, and point mutations in a control element or in the proto-oncogene itself. In addition, there exist genes, so called *tumor-suppressor genes*, whose normal products inhibit cell division. Any mutation that decreases the normal activity of these genes may contribute to the onset of cancer. Since more than one somatic mutation is generally needed to produce all the changes characteristic of a full-fledged cancer cell the incidence of cancer increases greatly with age.[64]

This is the view of carcinogenesis that is still prevailing in contemporary biomedical research. However, in recent years more and more critical voices can be heard. The target of their criticism is *genetic reductionism*, that is, the exclusive focus on the search for the genetic causes of cancer.[65]

---

[63] This revolutionist stance also finds expression in the titles of their papers, for example "Somatic Mutation Theory of Carcinogenesis: Why It Should Be Dropped and Replaced" (Soto/Sonnenschein 2000).

[64] The stance that underlies SMT can be condensed to the following slogan: "A gene is broken, fix the broken gene and cure disease" (Joyner and Pedersen 2011, 1018).

[65] The proponents of SMT more and more take into account also non-genetic factors, such as communication signals between cells and other extra-cellular factors (Hahn/Weinberg 2002; Malaterre 2007, 59). However, they do not regard changes at the tissue level as being important, as proponents of TOFT claim.

> The current carcinogenic paradigm is based on the assumption that all biological information is embedded in the DNA sequences insofar as any modification/mutation of the gene is thought to be linearly and automatically translated into a well-defined cellular abnormality. (Bizzarri, et al. 2008, 179)

> According to the reductionist view of cancer, expressed in myriads of molecular biology-based investigations, all the information necessary for a cell to transform itself into a neoplastic cell can be attributed to changes at the genomic level. (Grizzi, et al. 2006, 37)

The proponents of TOFT blame the advocacies of SMT for practicing genetic reductionism, too. They argue that the exclusive search for causes of carcinogenesis at the level of genes and of molecular components of cells leads researchers to overlook important causes at the level of tissues. According to TOFT, proliferation is the default stance of all cells and oncogenes act initially by disrupting the normal interactions between cells in the stroma and parenchyma of an organ (Soto/Sonnenschein 2005, 112). This is why the genesis of cancer is only adequately explained in a *non-reductive* fashion in terms of the disruption of tissue organization.[66]

The example of cancer research serves as an instructive illustration of the reductive character (more precisely, of the lower-level character) of genetic explanations. The phenomenon to be explained is the development of cancer, which is characterized as a tissue-level phenomenon. According to SMT, it is explained by appealing to certain mutations in genes and to the changes in the gene products that these mutations affect. That is, cancer is explained by referring only to such factors that are located on a lower level than the level of tissues, namely on the level of genes (or molecules in general).

One question remains. What is the relation between these genetic explanations and molecular explanations, which I named fundamental-level explanations? One major thesis in Sarkar's book on "Genetics and Reductionism" (1998) is that there is a significant difference between molecular explanations (which he calls "physical reductions"; 1998, 45) and genetic explanations (see Chapter III, Section 2.2). However, Sarkar focuses on genetic explanations as they can be found in classical genetics (i.e. on explanations of the origin of a trait by reference to a model of gene transmission). The genetic explanations I am concerned with here are different. They rely on the notion of a molecular gene, not on the concept of a Mendelian gene.[67] Accordingly, genes can be conceived as a subtype of molecules and genetic explanations as a *subtype of molecular explanations*. Hence, genetic explanations are fundamental-level explanations, too.[68]

---

[66] A few philosophers have picked up on this debate about the correct epistemological stance in cancer research and, for instance, examined the relationship between metaphysical presuppositions and scientific practices (Marcum 2005) or proposed a systemic approach to carcinogenesis (Malaterre 2007).

[67] This difference is explained for instance in Waters 1994.

[68] Genetic explanations can also be characterized as single-factor explanations (see Section 2.3.2). This is particularly true for genetic explanations that appeal to single genes or single kinds of genes (see fn. 71).

Not All Reductive Explanations Are Fundamental-level Explanations

The identification of reductive explanations with molecular explanations may be appropriate for specific scientific fields, in which the phenomena to be explained are located on a relatively low level itself (e.g. on the cellular level). However, philosophers making claims about reductive explanations or explanatory reductionism in general should refrain from equating reductive explanations with fundamental-level explanations. Otherwise they obscure the diversity of reductive explanation that is available in the biological sciences. Moreover, they run the risk of converting explanatory reductionism into a straw man that is too easy to reject. To explain biological phenomena reductively means to explain it exclusively by reference to lower-level objects, their properties and interactions. This *can*, but *need not* imply to explain them in molecular terms. The following example clarifies why.

Example: Muscle Contraction (Animal Physiology)

In the 1950s it was discovered that the contraction of muscle fibers[69] results from relative sliding between so called thick filaments (consisting mainly of myosin) and thin filaments (consisting mainly of actin) (Huxley/Niedergerke 1954; Huxley/Hanson 1954). Since then many molecular studies have been carried out to discover the molecular mechanism of muscle contraction. Today many of the details of what makes the filaments slide are known.

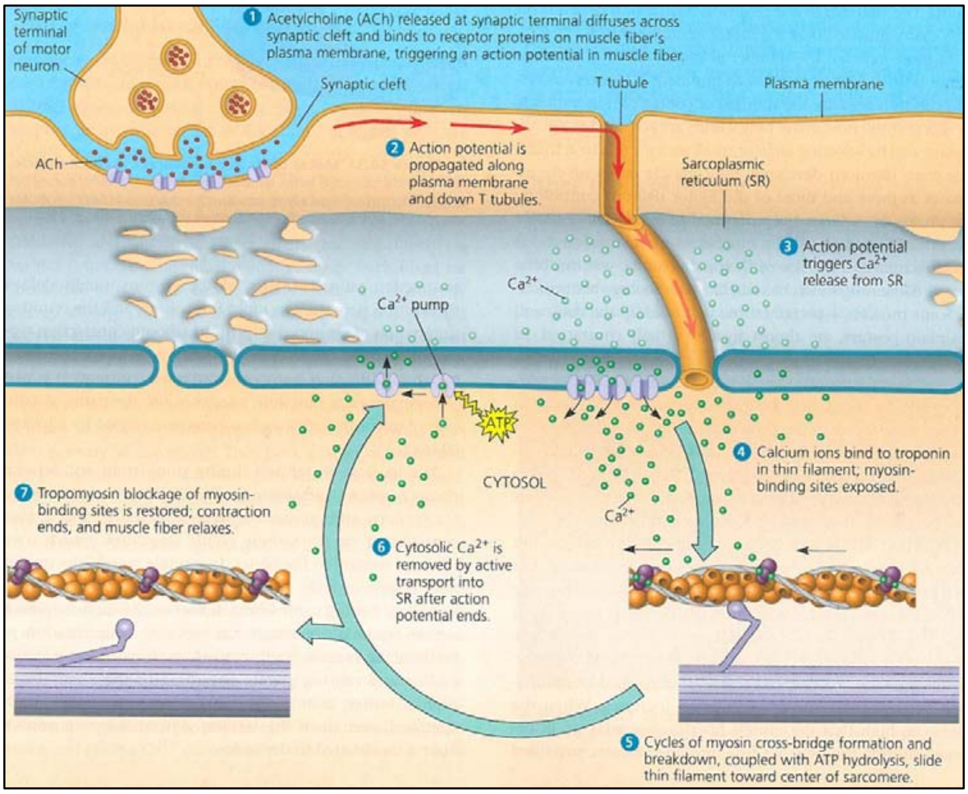

Figure V.16: Mechanism of Contraction in a Skeletal Muscle Fiber (Reece, et al. 2011, 1153)

---

[69] Muscle fibers originate during development from the fusion of several undifferentiated immature cells, called myoblasts, into long, cylindrical, multi-nucleated cells. Muscle fibers are composed of actin and myosin myofibrils repeated as a sarcomere, the basic functional unit of the muscle fiber. The term 'muscle' refers to multiple bundles of muscle fibers held together by connective tissue.

According to the *sliding-filament model* (which is depicted in the above figure) the process of muscle contraction starts with an incoming neuronal signal, that is, with an action potential that reaches the motor neuron terminal and causes a release of the neurotransmitter acetylcholine into the synaptic cleft. Acetylcholine binds to receptor proteins on the muscle fibers plasma membrane and triggers an action potential, which spreads through the muscle fiber's network of T-tubules, depolarizing the inner portion of the muscle fiber. The depolarization triggers the release of calcium ions (Ca^{2+}) from the sarcoplasmic reticulum into the cytosol. This is the point where more condensed descriptions of the mechanism of muscle contraction begin. In the absence of a neuronal stimulation the regulatory protein tropomyosin blocks the binding of myosin (thick filaments) to actin (thin filaments). This blocking is lifted as soon as calcium ions are released into the cytosol and bind to another set of regulatory proteins, the troponin complex, which removes the tropomyosin from the myosin binding sites on the actin filaments. Myosin (which has ADP and inorganic phosphate bound and is in the “cocked back” conformation) binds to the newly uncovered binding sites on the thin filaments. Myosin releases ADP and the inorganic phosphate, which is tightly coupled with a distortion of the myosin head. This power stroke causes that the thick and thin filaments slides past each other. ATP binds myosin and thereby allows it to release actin and be in the weak binding state. The myosin then hydrolyzes the ATP and uses the energy to move back into the “cocked back” conformation again. Myosin binds again to actin and the process of binding-sliding-releasing repeats as long as the neuronal signal is present (and ATP is available). When no new action potential triggers the release of calcium ions their concentration in the cytosol decreases (since the calcium ions are actively pumped back into the sarcoplasmic reticulum), tropomyosin blocks the myosin binding sites again and the muscle contraction ceases. In sum, the behavior of a muscle fiber to actively shorten (or tense) itself is explained by referring to various parts of the muscle fiber (myosin, actin, calcium ions, tropomyosin, ATP, etc.), which primarily are *molecular parts* (apart from e.g. the sarcoplasmic reticulum, a cell organelle), and the various interactions (binding, blocking, releasing, triggering, etc.) between them that bring about the sliding of the thick and thin filaments past each other.

Against this background one might wonder whether one needs to go down to the molecular level in order to explain the phenomenon of muscle contraction. In fact, there is an alternative (but not incompatible) explanation available. This explanation focuses on the structure of a muscle fiber and its basic functional unit, the sarcomere.

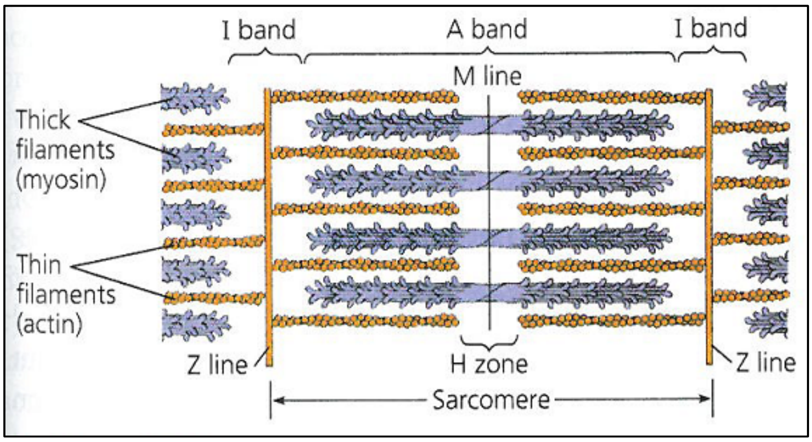

Figure V.17: The Sarcomere of a Skeletal Muscle (Campbell, et al. 2005, 1067)

The contraction of a muscle fiber is explained by showing how the sliding of the thick and the thin filaments past each other gives rise to the shortening of the I-bands and H-zone until they, in the case of full contraction, disappear. Moreover, the explanation describes how the Z lines move towards each other, whereas the A bands remain the same length. All in all, the explanation reveals how the entire muscle fiber is shortened.

This example demonstrates that for a single biological phenomenon there can be two adequate explanations available, which *both* are to be characterized as reductive because both possess a lower-level character. The phenomenon to be explained is the behavior of a muscle fiber (which is a multinuclear cell). It can be explained, on the one hand, in terms of the interactions between specific molecules (most importantly, between actin, myosin, tropomyosin, and ATP), and on the other hand, by reference to the thick and thin filaments as well as to how certain regions of the sarcomere change their sizes. Both explanations are reductive because both explanations appeal only to *lower-level* factors. Actin, myosin, tropomyosin, and ATP are genuine parts of the muscle fiber as well as thick filaments, thin filaments, and sarcomeres. There is no reason why the term reductive explanation should be restricted to the former, the molecular explanation, and not be applied to the latter, non-molecular explanation, too.

Especially when it comes to disciplines such as ecology and evolutionary biology it seems misleading to limit the notion of a reductive explanation to fundamental-level explanations. Recall, for instance, the reductive explanation of the food web structure of a community in ecology. So-called individual-based models (IBMs) explain the properties or behavior of a community in a reductive way because they refer merely to individual organisms and to the interactions between them.

In philosophy, many authors who dispute about whether explanatory reductionism is true or not disregard the fact that to explain a phenomenon reductively is not identical with explaining it in molecular or physicochemical terms (e.g. Fodor 1974, Nagel 1952, de Vreese, et al. 2010, Rosenberg 2006, etc.). Of course, fundamental-level explanations are an important subtype of reductive explanations, but they are not the only important kinds of reductive explanations that can be found in biological practice. Identifying reductive explanation with molecular explanation or genetic explanation overlooks the diversity of reductive reasoning that is present in contemporary biological practice. An additional problem is that in some contexts it turns antireductionism almost into a truism. For instance, today almost everyone agrees that we cannot explain complex phenotypic traits by appealing exclusively to genetic causes. However, the far more interesting question is whether we can explain them by reference only to lower-level factors. That is, do we need to take into account the context of an organism, too, or is it adequate to explain particular phenotypic traits in terms of the parts of an organism and other lower-level factors. Since far more factors count as lower-level than the genes of an organism, the identification of reductive explanation with genetic or molecular explanation hides this interesting question.

2.3.2 Single-factor Explanation

In the biological literature there exists another way of speaking about reductionism, which points to another subtype of reductive explanation. Regard for example the following statements that stem from molecular biology, cancer research, and ecology:

> Reductionists favour causal explanations that give undue explanatory weight to a *single factor*. (van Regenmortel 2004b, 1018; my emphasis)
>
> These new approaches clearly highlighted the shortcomings of the reductionistic and mechanistic thought, according to which even more complex biochemical pathways… are governed by a *single pacemaker gene or enzyme*. (Bizzarri, et al. 2008, 180; my emphasis)
>
> [I]t is reductionistic to explain body size with a *single factor* and holistic to explain it with a number of factors that interact with each other. […] In this fashion the term holistic is often used merely to say that more variables must be considered to understand the object of study. […] [E]cology and evolution abound with single-factor explanations that turn out to be inadequate and yield to multi-factor explanations. […] There is a very important class of people who disagree with this assessment, however: the people who advance the single-factor models. They usually claim that their single factor alone suffices to explain the behavior of the whole, and resist the idea that their factor might be context dependent or interact with numerous equally important factors. (Wilson 1988, 270, my emphasis)

According to these quotes, the reductionistic procedure can be characterized as studying a phenomenon by taking into account only a single (causal) factor, even if this phenomenon is influenced by several different factors. To explain a phenomenon reductively means to cite only a *single (causal) factor* and to ignore other factors, which might be relevant to the explanandum phenomenon, too. In other words, a reductive explanation "provides only part of the causal story" (Byerly 2003, 336). I call these explanations *single-factor explanations*.[70]

---

[70] One may wonder why this type of reductive explanation is discussed in this section, that is, under the label of the lower-level character of reductive explanations, and not, for instance, in the context of the internal

It is important to note that single-factor explanations rarely occur in biological research practice.[71] Instead, this notion of reductive explanation is present almost exclusively in contexts in which reductionism is criticized (e.g. Ahn, et al. 2006b, 1; Bizzarri, et al. 2008; Levenstein 2009, 709; Wilson 1988; van Regenmortel 2004a; Byerly 2003) or in which philosophers discussing reductionism want to emphasize that this is *not* what they mean by reductionism (e.g. Rosenberg 2006; Weber 1998, 293). Rosenberg, for example, states:

> Most of the time, reductionism is a term of abuse. It is employed indifferently to identify two different explanatory strategies in science. One of them is clearly a mistaken strategy, and rightly criticized. This is the temptation to simply ignore causal variables in explaining an outcome. […] [I]t would be wrongly reductionistic to trace the causes of a chocolate Labrador's coat color to its genes. Not even all of its genes alone will produce pigment or pattern by themselves. They may be necessary, but they are not sufficient for coat color. […] The sort of reductionism that simply neglects causally necessary factors is one that doesn't require too much refutation. But the sort of reductionism found to be threatening in biology is quite a different doctrine. (2006, 11f)

I agree with Rosenberg on the point that a reductionism which is grounded merely in single-factor explanations is not very threatening or challenging. It is too obvious that only very few phenomena are adequately explained by single-factor explanations. What is more, Rosenberg is right in emphasizing that this does not make the question of reductionism redundant. The reason is that biological explanations *need not* be single-factor explanations to exhibit a reductive character (just as they need not be single-level explanations). For an explanation to be reductive it suffices for it to be a lower-level explanation, that is, to explain a phenomenon by referring only to lower-level factors.

3 Internal Factors

In this section I present a second feature that is typical for reductive explanations in the biological sciences, namely that they explain the behavior of a system by focusing on those *factors that are internal to the system* (i.e. that are genuine parts of the system). In other words, my central claim will be that most reductive explanations in biology are part-whole explanations. The term 'most' expresses an important result of my analysis, namely that the internal character is typical for many reductive explanations in the biological science, but not for all. In other words, it is just a typical feature, but *not* a *necessary condition* for the reductivity of an explanation.

---

character of reductive explanations (Section 3). I admit that one could allocate this subtype differently. However, I think that it suits this section well because the single factor is chosen among the lower-level factors.

[71] Examples are genetic explanations that refer to the one gene that is taken to cause a particular phenotypic trait. Explanations of this kind are often reported by such statements as: "The gene for x (e.g. aggression, obesity, crib death, intelligence, etc.) has been found."

In Section 3.1 I will introduce the internal character of reductive explanations by analyzing how it is discussed in the biological literature on reductionism and by examining an instructive example of how an explanation fails to exhibit an internal character. Subsequently, I highlight the difference between the internal character and the lower-level character of an explanation (Section 3.2). In Section 3.3 I analyze in more detail what it means that an explanation focuses on factors that are internal to the system of interest. I will argue that this amounts to either ignoring environmental factors altogether or simplifying them in a specific manner.

3.1 Focusing on Internal Factors

In the biological literature a second feature of reductive explanations is mentioned – far from as frequently as the lower-level character, but nevertheless frequently enough to be important to my analysis. This feature is implicit in how biologists specify the reductionist methodology. When biologists discuss the limits of reductionism, they assume that a reductionist embraces not only the "search… downward in the hierarchy of organic organization" but also the "search *inward* [i.e., into the system]" (Lidicker 1988, 278; my emphasis). In other words, "advocacy of a reductionist approach coincides with emphasizing *internal*, rather than external, factors." (Schoener 1986, 102; my emphasis; see also Mikkelson forthcoming, 4) By contrast, non-reductionistic research (also called 'systemic' or 'holistic') is associated with "putting things in a context" (Grizzi/Chiriva-Internati 2005, 28) and "look[ing] outwardly from the boundaries of the phenomenon under study" (Lidicker 1988, 280). The opposition that is portrayed in these statements is the following: when biologists investigate the behavior of a certain system with a reductionist methodology they focus on factors that are *internal* to the system and ignore or simplify factors that belong to the environment or context of the system (i.e. *external* factors). In contrast, a non-reductionistic proceeding also takes into account the context of the system and, for instance, explores how changes in the environmental conditions affect changes in the interactions between the system's parts and how this in turn influences whether the system displays the behavior in question or not. This reductionist research strategy can also be characterized as shielding the system from its environment and treating biological systems as if they were closed or semi-closed systems. Such a procedure can be beneficial because it makes the investigation especially of complex systems more manageable and, thus, the construction of explanations of the system's behavior easier.[72] But since biological systems generally are open systems as well as systems that evolved under the influence of specific environmental conditions, in most cases the environment is crucial for the functioning of these systems and cannot be as easily ignored and simplified as the reductionist might

---

[72] Bechtel and Richardson put a lot of effort in depicting these "psychological constraints" for developing explanations (see, e.g., 2010, 234-243).

wish.[73] The more important the environment is the more is the adequacy of the reductionistic methodology called into question.

From this it follows that reductive explanations in the biological science exhibit a second characteristic: they explain the behavior of a system by focusing on those entities and interactions that are *internal* to this system (this is called the internal character of a reductive explanation). I understand the term 'internal' in the following sense: any factor is internal to a particular system S that is a *genuine part* of S. In Section 1.2 I characterized the concept of a genuine part of S as referring to all those interacting entities that, first, contribute to bringing about at least one of the behaviors/functions that S typically exhibits and that, second, are located inside the spatial boundary of S. Because everything that is internal to a system is at the same time a genuine part of the system (and the other way round), the set of all explanations with an internal character is identical to the set of all *part-whole explanations* (recall that this is not true for lower-level explanations; more on this in Section 3.2). The internal character of reductive explanations (just as their lower-level character) reveals that reductive explanations possess a *clear direction*: to explain a phenomenon reductively implies to look inside the system in question, not outside of it (likewise, it implies to look at lower levels of organization than the level of the system in question, not on the same or on higher levels).

The fact that reductive explanations exhibit not only a lower-level character, but also an internal character (more on the difference between the two in Section 3.2) has not been recognized in the philosophical literature for a long time. If the internal character were mentioned at all, it has been lumped together with the lower-level character (e.g. Wimsatt 1976a; Sarkar 1998, 2005). There are two notable exceptions. In more recent years Wimsatt has developed a list of "common reductionist heuristics" (2006a, 467; 2007, Appendix B), in which he also specifies reductionist heuristics that correspond to or are connected with the internal character of reductive explanations.[74] In addition, Hüttemann and Love (2011) point out that the internal character and the lower-level character – they call them "intrinsicality"[75] and "fundamentality" (2011, 527) – are two distinct "aspects" (2011, 523) of reductive explanations (see Chapter III, Section 2.3).

---

[73] Exceptions are for instance *homeostatic systems*, which possess the capacity of self-regulation, i.e. of sustaining the internal conditions under a certain range of variations in the environmental conditions. The most common kind of homeostasis is the regulation of the body temperature that can be found in endothermic animals like birds and mammals (see Reece, et al. 2011, 906-914).

[74] Consider the following reductionistic heuristics: "(3) *Interface determinism*: […] *black-world perspectivalism* – all that matters about the environment is what comes in across system boundaries… […] (5) *Modelling localization*: look for an intra-systemic mechanism to explain a systemic property rather than an inter-systemic one. […] (6) *Contextual simplification*:… simplify environment before simplifying system. […] (7) *Generalization*:… focus on generalizing or elaborating the internal structure, at the cost of ignoring generalizations or elaborations of the environmental structure." (2006a, 468f)

[75] It might be that the term 'intrinsic' is more suitable than the term 'internal' in the respect that the former might be spelled out in a non-spatial way, too, and thus captures also more rare types of reductive explanations (e.g. structural explanations). However, I stick to the term 'internal' because of two reasons. On

Just as an explanation can fail to possess a lower-level character it can fail to exhibit an internal character, too. This is the case when certain contextual factors are so crucial to the behavior of the considered system that they cannot be ignored or simplified, but rather need to be included in the explanation. The following example of the explanation of protein folding illustrates such a failure of the internal character of an explanation (and of its lower-level character, too).

Example: Protein Folding (Molecular Biology)

Protein folding is the process by which functioning proteins arise from linear chains of amino acids. Proteins are composed of many different amino acids that are linked into a linear chain, called polypeptide. Since the specific activities of a protein result from its intricate three-dimensional configuration, the linear chain of amino acids (the *primary structure* of the protein) needs to be folded after it has been produced. In protein folding different steps[76] can be distinguished that correspond to different levels of protein structure: The *secondary structure* originates when hydrogen bounds are formed between the repeating constituents of the polypeptide backbone. The resulting types of structures are either delicate coils, known as ‘α helices’, or strands of polypeptide chains lying side by side, called ‘β pleated sheets’. Superimposed on the patterns of secondary structure is a protein’s *tertiary structure*, which results from interactions between the various side chains of the amino acids. Depending on their properties, the side chains form hydrogen bounds, disulfide bridges, ionic bounds, or become involved in hydrophobic or van der Waals interactions with neighboring side chains. If a protein is composed of only one polypeptide chain the tertiary structure is its final configuration. Proteins that consist of two or more polypeptide chains (e.g. hemoglobin) form a *quaternary structure*, which is the overall protein structure that results from the aggregation of its polypeptide subunits.

If this were the complete story to be told the folding of a protein could be explained by the chemical interactions between its constituents (i.e. the amino acids) alone.[77] Such an explanation would be reductive in two respects: first, it would be reductive because it would exclusively refer to interacting entities (i.e. amino acids forming different kinds of bounds between each other) that are not only located on a *lower level* than the phenomenon to be explained (i.e. protein folding) but on the fundamental level of molecules. Second, it would be reductive because it would appeal only to factors (i.e. amino acids) that are *internal* to the system whose behavior is to be explained (i.e. polypeptide or protein).

---

the one hand, the biological literature shows that the notions of reductionism and of a reductive explanation are closely linked to the idea that one considers only entities and interactions that are spatially located inside the system. On the other hand, the term ‘intrinsic’ has the drawback that it confounds the internal character with the third characteristic of reductive explanations, which is introduced in Section 4.

[76] These “steps” are distinguished for heuristic reasons. In reality, they do not occur in such a neat, clearly successive fashion as suggested here (see e.g. Dobson 2003).

[77] This thesis is sometimes referred to as (one reading of) the “linear sequence hypothesis” (Hüttemann/Love 2011, 17), in short “LSH”.

However, during the last two decades researchers have figured out that this is *not* the complete story that is to be told. Environmental factors do play a crucial role for the folding of most proteins. The consequence is that they cannot be just ignored and, perhaps, also not simplified in explaining the folding of proteins. One class of contextual factors includes a certain temperature, pH value, and salt concentration in the cytosol, also referred to as the 'normal physiological milieu'. These factors are necessary since inadequate temperatures, pH values, and salt concentrations change the properties of the side chains of the amino acids, which forecloses the correct folding of a polypeptide chain into a functioning protein. But although these factors are important I see no reason why they cannot be conceptualized as being mere background conditions that remain constant over time and, thus, must not be explicitly mentioned in an explanation of protein folding.

Things look differently when we are concerned with the fact that proteins need the assistance of molecular *chaperones* in order to fold. Some chaperones interact with nascent chains as they emerge from the ribosome, whereas others are involved in guiding later stages of the folding process. The function of chaperones is not to increase the rate of individual steps in protein folding. Rather, they increase the efficiency of the folding by insulating them from the crowded intracellular milieu, which prevents misfolding and aggregation (Hartl/Hayer-Hartl 2002; Dobson 2003; Frydman 2001).

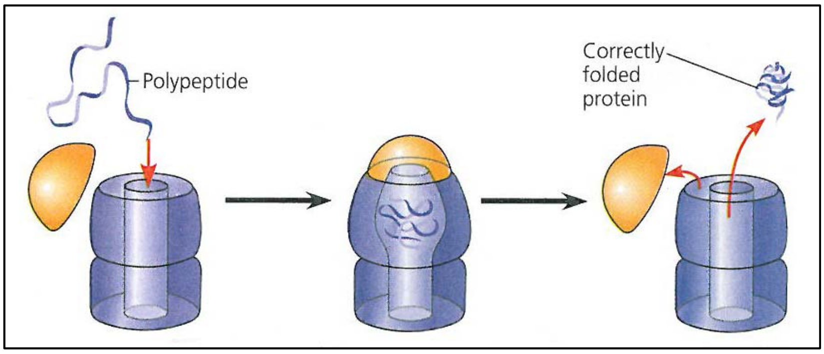

Figure V.18: Steps of Chaperonin Action (Reece, et al. 2011, 131)

The high degree of attention biologists devote to the influence molecular chaperones have on the process of protein folding suggests that chaperones are more than just necessary background conditions. Instead, they are an essential part of the explanation of how linear chains of amino acids fold into functioning proteins with a characteristic three-dimensional structure. Chaperones are also more than mere input conditions since different chaperones provide different kinds of assistance in earlier and latter stages of protein folding; and even one chaperone protein can impact the folding process differently to different times (see Frydman 2001). Thus, the reductionist strategy of excluding chaperones from the explanation of protein folding or of simplifying them as mere background or input conditions (see Section 3.3) would lead to an inadequate, incomplete explanation of protein folding. The only way to develop an adequate explanation of protein folding is to include chaperones as explanatorily relevant factors in the explanation (as biologists actually do). Such an explanation is *non-reductive* because it refers to environmental factors (i.e. chaperones), which are located on a higher level (respectively on the same level) as the explanandum phenomenon (i.e. the polypeptide or protein whose folding is explained).

In sum, the example of protein folding presents a clear case in which the reductionist strategy of focusing exclusively on factors that are internal to the system under study reaches its limits. In cases like this the environment of a system can only be ignored or simplified at the cost of the inadequacy of explanation.

3.2 Distinguishing the Internal from the Lower-level Character

One might raise the objection that the internal character is not sufficiently different from the lower-level character, which is why one can refrain from drawing this distinction. First of all, what is true about this objection is that the internal character of reductive explanations is *closely related* to their lower-level character. Any factor that is internal to a particular system S is a genuine part of S and, thus, automatically located on a lower level than S. Accordingly, if an explanation exhibits an internal character, it displays a lower-level character, too. To put it another way, the set of all explanations with an internal character is a subset of the set of all lower-level explanations.

The latter, however, implies that there exist lower-level explanations that do not possess an internal character. This is exactly the reason why 'exhibiting an internal character' must be distinguished from 'exhibiting a lower-level character'. As I have already mentioned, it is possible that explanations appeal to environmental factors and, nevertheless, are lower-level explanations. This requires that they refer only to those entities and interactions of the environment that are located on a lower level than the system in question.[78] Recall, for instance, the explanation of muscle contraction. This explanation involves referring to an incoming neuronal signal – an entity which belongs to the context of the muscle fiber because it is located outside its spatial boundary (i.e. its cell membrane). However, the incoming neuronal signal can be represented as an action potential that reaches the motor neuron terminal and causes a release of the neurotransmitter acetylcholine into the synaptic cleft. Described in that way, the incoming neuronal signal is located on a lower level than the muscle fiber. Hence, the explanation of muscle contraction does not possess an internal character because it appeals to a factor that is not a genuine part of the muscle fiber, but that rather belongs to the context of the muscle fiber. But yet, this explanation is a lower-level explanation since it refers only to lower-level factors.

---

[78] This presupposes an account of levels of organization, according to which levels are not only locally defined, but according to which levels are *global levels*. Only the notion of a global level allows one to make claims about the level on which contextual factors are located (see Section 1.3.3, 1.3.4, and 2.2.1).

Reductive explanations of this kind (i.e. which are lower-level but not internal) occupy center stage in the reductionist's response to what has been called the "context objection to reduction" (Delehanty 2005, 715). According to the context objection, explanatory reductionism is problematic since many biological phenomena can only be adequately explained by taking into account their context, too. A reductionist can counter that reductionism can be preserved by reducing also the context. The phrase 'reducing the context' means that it is spelled out in lower-level terms. This suggests that there exist explanations, which do not only refer to internal factors, but which are still reductive because the appeal only to lower-level factors.

To conclude, since the category of lower-level factors is broader than the category of internal factors, the set of all part-whole explanations/explanations with an internal character constitutes only a *subset* of the set of all lower-level explanations. In line with this, the internal character should be regarded as a *typical feature* that many reductive explanations in biology exhibit, but *not* as a *necessary condition* for an explanation to be reductive. That is, unlike the lower-level character the internal character is not a necessary condition for an explanation to be reductive. Accordingly, to reconstruct the internal character as a second major feature of reductive explanations does not widen the concept of explanatory reduction in the sense that more cases of reductive explanations are captured than without recognizing it. Instead, this feature adds complexity and makes explicit a distinction that was not apparent before.

3.3 Simplifying the Environment

So far I have specified the internal character of reductive explanations by saying that reductive explanations explain the behavior of a system by focusing on those entities and interactions that are internal to the system. In this section I spell out in more detail what the expression 'by focusing' means. In reductive explanations are environmental factors ignored, that is, excluded from the explanation altogether? If they are not, in which form are they mentioned in a reductive explanation so that the explanation is still taken to focus merely on internal factors?

The biological literature does not provide many indications of how these questions are to be answered.[79] However, if we analyze paradigmatic cases of reductive explanations

---

[79] I only found two statements of biologists that are relevant to these questions. However, they provide only limited insights. For instance, Greenspan characterizes reductionism (in molecular genetics) as "the neat view of biological systems made up of dedicated components… and in which particular *starting conditions* give rise to uniquely predictable responses" (2001, 386; my emphasis). Furthermore, Levins and Lewontin assume that with respect to ecology "reductionism takes the form of regarding each species as a separate element existing in an environment that consists of the physical world and of other species. The interaction of the species and its environment is *unidirectional*: the species experiences, reacts to, and evolves in response to its environment. The reciprocal phenomenon, the reaction and evolution of the environment in response to the species, is put aside." (1980, 49)

in biology we discern that the internal character of reductive explanations does not imply that environmental factors *must* be excluded *completely* from the explanation. Instead, it can also mean that a reductive explanation appeals to environmental factors, but that these environmental factors are *simplified* to a great extent. My analysis yields that two kinds of simplification are of particular importance in this context:

First, environmental factors can be conceptualized as being mere *background conditions*, which are homogeneous or remain constant over time, and are thus only implicitly included in the explanation. An example is the reductive explanation of protein folding which does not explicitly mention external factors as a certain temperature, pH-value, and salt-concentration but which, nevertheless, implicitly includes these factors as background conditions necessary for the folding to occur.

Second, contextual factors can be simplified as being pure *input conditions* (or starting conditions)[80], which need to be satisfied at the beginning of the process that brings about the explanandum phenomenon “automatically” (i.e. without further help of other environmental factors). Consider again the explanation of muscle contraction. In a reductive explanation of this phenomenon only a single environmental factor, namely the incoming neuronal signal, is mentioned. In addition, this factor is simplified since the only thing that matters is that it is present and that it induces the entire process of muscle contraction. The exact nature of this signal, how and where it originates and how it is transmitted to the muscle fiber, is irrelevant to the explanation.

The following example illustrates and further supports my thesis that in reductive explanations contextual factors are either ignored altogether or simplified as being mere background or input conditions.

Example: The PI 3-Kinase/Akt Signaling Pathway (Molecular Biology)

In the most well-known textbook for molecular biology, “the Alberts” (Alberts, et al. 2008), the chapter on “Mechanisms of Cell Communication” is introduced in the following way: “To make a multicellular organism, cells must communicate, just as humans must communicate if they are to organize themselves into a complex society.” (2008, 879) Cell-cell communication is mediated mainly by *extracellular signal molecules*, some of which operate over long distances; others affect only cells close-by. The way a cell responds to the different combinations of signal molecules it can be exposed to in its environment depends on various factors, in particular on the cell-surface receptor proteins a cell exhibits. Various processes in a cell are regulated in this manner, for instance cell survival, growth, division, differentiation, and dead (apoptosis). One way in which cell survival is promoted by signaling through the phosphoinositide (PI) 3-kinase/Akt pathway is presented here:

---

[80] Machamer, et al. speak about start or set-up conditions, too (2000, 11). However, there is an important difference between our positions: they regard these conditions to be parts of a certain mechanism, whereas I think that (in most cases) it is more appropriate to interpret them as parts of the context of a system, not as genuine parts of the system itself.

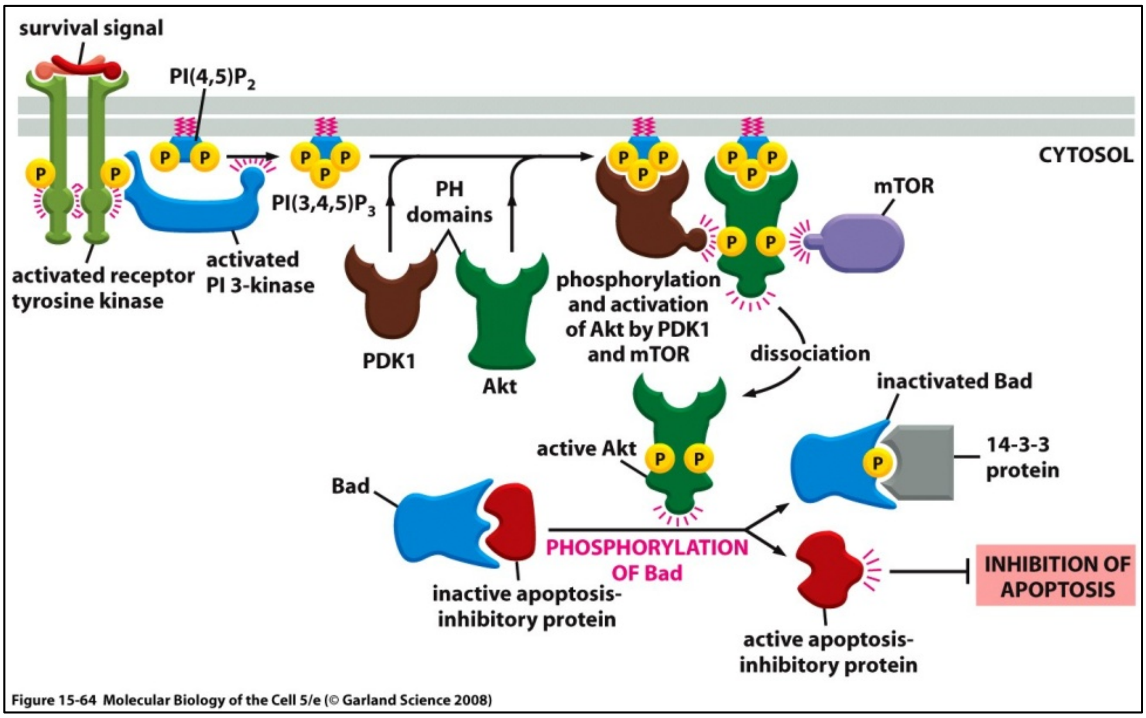

Figure V.19: How the PI 3-kinase/Akt Pathway Stimulates Cell Survival (Alberts, et al. 2008, 934)

The explanation of how the survival of a cell is promoted (the explanadum phenomenon) starts with the binding of an extracellular survival signal (e.g. a member of the insulin-like growth factor (IGF) family) to RTK, which is thereby activated and, in turn, recruits and activates PI 3-kinase. The PI 3-kinase produces PI(3,4,5)P$_3$, which serves as a docking site for Akt and PDK1. Akt is phosphorylated by a third kinase (usually mTOR), which alters the conformation of Akt so that is can be phosphorylated at a second place by PDK1. This second phosphorylation activates Akt, which then dissociates from the plasma membrane and phosphorylates several target proteins, including Bad. When Bad is phosphorylated it releases one or more inhibitory proteins, which now can block apoptosis and thereby promote cell survival. (cf. Alberts, et al. 2008, Chapter 15; Kennedy, et al. 1997; Shimamura, et al. 2003) This specific process is only one small part of a whole network of signal transduction pathways involved in the regulation of apoptosis and cell survival:

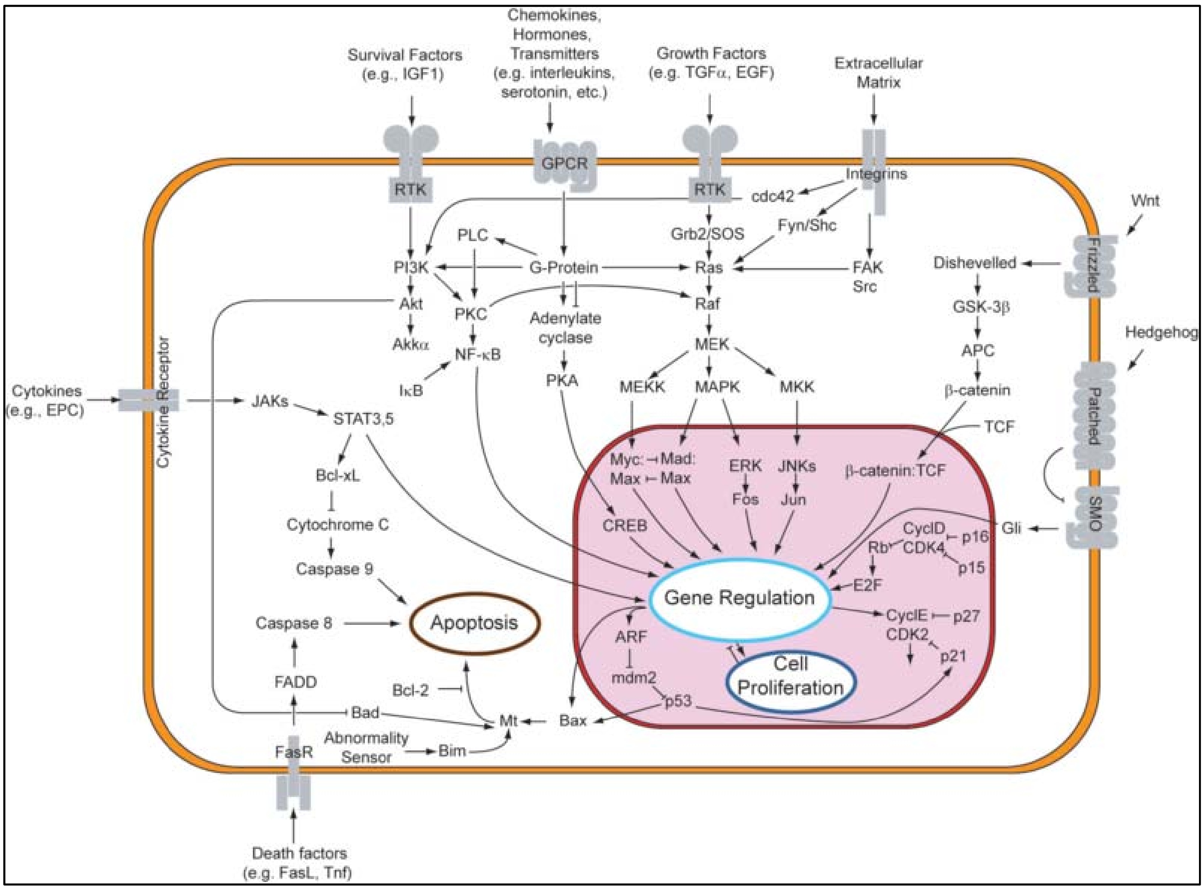

Figure V.20: Overview of Signal Transduction Pathways Involved in Apoptosis

In what way does this example provide further evidence for my thesis that in reductive explanations environmental factors are frequently simplified as being mere input conditions or conceptualized as background conditions? The phenomenon to be explained is the behavior of an animal cell to survive. The boundary of the system in question (i.e. the cell), which separates it from its environment, is the cell membrane (in Figure V.19 colored gray, in Figure V.20 orange). The explanation of how a cell survives focuses on the processes that go on inside the cell. This is why Figure V.19 almost exclusively depicts entities and interactions that are internal to the cell (i.e. located inside the cell membrane). The only external factor that is described is the extracellular survival signal molecule and it's binding to RTK[81]. However, even if the explanation appeals to the survival factor, it seems to me that the explanatory role it plays is different from those of the internal factors (like PI 3-kinase, Akt, Bad, etc.). More specifically, the survival factor is represented as being a pure *input condition* that marks the starting time of the intracellular process that brings about the survival of the cell (i.e. the phenomenon to be explained). Granted, the presence of survival signal molecules is causally necessary for the explanandum phenomenon to occur. Without them the process would not run. However, neither the exact character of this external factor

---

[81] RTK is a transmembrane receptor protein. As such it is a part of the cell membrane and, hence, an internal factor.

is specified (e.g. the explanation only says "for example members of the insulin-like growth factor family"), nor is it stated what its origin is, why and how it has been transported to this particular cell. The only thing that matters is that a survival signal molecule is present. I take this to indicate that the presence of the environmental factor (i.e. the survival signal molecule) is simplified as being a mere input condition.[82]

Other parts of the environment of the cell (e.g. extracellular inhibitory molecules, other extracellular signal molecules, neighboring cells, gap junctions between cells, the tissue or organ the cell is localized in, etc.) are ignored altogether. Some of these ignored contextual factors are simply causally irrelevant to the specific phenomenon that is to be explained. Others are not causally irrelevant but are, for reasons of simplicity, assumed to be absent or constant. In my words, they are classified as mere background conditions and, thus, ignored in the explanation. Examples of this kind of factor include extracellular inhibitory molecules or other signal molecules, which interfere with survival signal molecules or neutralize their effect.

In sum, many reductive explanations in the biological sciences exhibit an internal character. This amount either to ignoring environmental factors in the explanation of a particular behavior of a system altogether, or to including selected parts of the system's environment, but simplifying them. The context of a system can be simplified in two major ways: on the one hand, environmental factors can be simplified as background conditions, which are treated as being homogeneous or remaining constant over time. On the other hand, they can be simplified as mere input conditions, which need to be satisfied only at the beginning of the process that brings about the explanandum phenomenon "automatically" and whose specific character and origin is regarded as irrelevant.

4 Parts in Isolation

In biological papers on reductionism a third characteristic of reductive explanations is implicitly mentioned. Philosophers have not paid much attention to it yet but the current vivid discussions in the biological sciences demonstrate that they should. A widespread argument that biologists put forward against reductionism goes as follows: reductionism is deficient because the behavior especially of complex systems cannot be understood (and explained) by dissecting the system into parts and by investigating the "parts in isolation" (Chong/Ray 2002, 1661), without "putting… [them] together again" (Keller 2005, 4) and understanding "how all these things [i.e. the parts] are integrated" (Service 1999, 81) into the system as a whole. My central goal in this section is to clarify what exactly it means to

---

[82] In accordance with my thesis, Wimsatt identifies a certain reductionistic heuristic, which he calls "*black-world perspectivalism*". He characterizes this reductionistic strategy as the assumption that "all that matters about the environment is what comes in across system boundaries" and that "the order of study is from a system with its input-output relations to its subsystems with theirs, and so on." (Wimsatt 2006a, 468)

study the parts of a system "in isolation" and to disregard how the parts are put together and integrated into a system.[83] What is more, I show in which sense these statements point to a third feature that is crucial for reductive explanations in biology, namely that they *appeal only* to the *system's parts in isolation*. An important result of my analysis will be that this third characteristic constitutes a *necessary condition* for a biological explanation to exhibit a reductive character.

After reviewing the biological literature on the merits and, in particular, the limits of reductionism (Section 4.1.1) I argue for how the phrase 'parts in isolation' is to be best understood (Section 4.1.2). My central claim will be that reductive explanations refer only to those relational properties of the parts and to those kinds of interactions that can be discovered by studying the parts in other contexts that *in situ*. In Section 4.2 I relate this thesis to the distinction of different kinds of systems (respectively of decomposability) that I introduced in Section 1.2.4. I examine whether the reductionistic methodology can be characterized as treating biological systems as if they were aggregative systems or as if they were nearly decomposable systems.

Before I start, let me add a general remark. It is important to overcome the temptation to interpret the notion of a reductive explanation in such a way that it promotes either the defense of reductionism or the criticism of reductionism (depending on the position one favors). Reductionists tend to endorse a broad concept of reductive explanation, according to which reductive explanations can appeal to almost all kinds of explanatory relevant factors, let it be the parts of a system, their intrinsic properties, their relational properties (including functional properties), the interactions between them, or the way they are spatially and temporally organized (e.g. Schaffner 1969, 345; Wimsatt/Sarkar 2006). Some authors even state that explanations referring to contextual factors or to higher-level factors can still be characterized as reductive (e.g. Bechtel 2006, 40-44; Bechtel/Hamilton 2007, 405-411; see also Section 5). By contrast, proponents of antireductionism commonly adopt a notion of reductive explanation that is much narrower. Some argue that reductive explanations do not include (all of) the organization of and the interactions between the parts of a system (e.g. Dupré 1993, Chapter II; 2009). Others even identify reductive explanations with explanations, in which a biological system is treated as being no more than the sum of its parts (i.e. as being an aggregative system). Applying a broad concept of reductive explanation promotes explanatory reductionism because it entails that reductive explanations are quite ubiquitous in biological practice. As opposed to

---

[83] It is not the case that no philosopher has paid attention to this issue so far. For instance, in the chapter on "Mechanisms and its Alternatives" in his book "The Mind and its Place in Nature" (1925) Broad makes the following claim: "It is clear that in *no* case could the behavior of a whole composed of certain constituents be predicted *merely* from a knowledge of the properties of these constituents, *taken separately*" (1925, 63; my emphasis). In his paper on "The Watson-Crick model and Reductionism" (1969) Schaffner addresses the same issue: "given an organism composed out of chemical constituents, the present behavior of that organism is a function of the constituents as they are characterisable *in isolation* plus the… causal inter-structure of the chemical constituents" (1969, 346; my emphasis).

this, endorsing a narrow notion of reductive explanation implies that reductive explanations are rarely adequate and thus not widely used in the biological sciences. My aim is to develop an account of reductive explanation in biology that is as unaffected by these philosophical disputes about explanatory reductionism as possible. That is, I seek to capture what reductive explanation in contemporary biological research practice is, without caring about whether this notion promotes a particular version or explanatory reductionism or antireductionism.[84]

4.1 Isolating Parts from their Original System

What does it mean when biologists accuse reductionists of studying and explaining the behavior of a system by considering only the system's parts in isolation? Let us start with examining how biologists specify this notion when they discuss the merits and "limits of reductionism" (Ahn, et al. 2006a, 709) in biological science.

4.1.1 Current Discussions about the Limits of Reductionism

In general, the worry is that reductionistic research runs the risk of *underestimating* the *complexity* of biological systems.[85] This danger arises because to investigate the behavior of a system in a reductive manner implies to decompose the system into its parts and to try to achieve an understanding of the system as a whole by studying the behavior or properties of the parts *in isolation*. The following quotes exemplify this thesis:[86]

> The reductionist myth of simplicity leads its advocates to isolate parts as completely as possible and study these *isolated parts*. (Levins/Lewontin 1980, 76; my emphasis)
>
> Reductionists analyze a larger system by breaking it down into pieces and determining the connections between the parts. They assume that *isolated molecules* and their structure have sufficient explanatory power to provide an understanding of the whole system. (van Regenmortel 2004b, 1016; my emphasis)
>
> In the past decades, an enormous amount of precious information has been collected about molecular and genetic characteristics of cancer. This knowledge is mainly based on a reductionistic approach... The behavior of complex physiological processes cannot be understood simply by knowing how the parts work *in isolation*. (Bizzarri, et al. 2008, 173; my emphasis)

---

[84] I admit that the unbiased character of my account might be threatened a bit by my strategy to take into account also the way how biologists dispute about reductionism. It might be argued that their notion of reductive explanation is affected by their wish to defend or to criticize reductionism, too. In fact, this may be applicable to some of their statements. However, I think that I managed to identify these cases and to treat them with special care. This will become apparent in the following sections.

[85] "The reductionist approach… underestimates this complexity…" (van Regenmortel 2004b, 1016). "[R]eductionist thinking… strives to achieve a representation of reality free of complexity." (Soto/ Sonnenschein 2010, 369)

[86] See also Greenspan 2001, 386; Kitano 2002, 1662; Chong/Ray 2002, 1661; and Strange 2005, 968.

This gives rise to the question what is wrong with the reductionistic strategy to explore the parts of a biological system in isolation. A system biologist provides the following answer:

> [G]ene sequencing and other techniques will soon have isolated all the cell's individual parts and spelled out their isolated functions. Now, it is time to move beyond reductionism. […] Now we need to know how all these things are *integrated*. (Service 1999, 81; my emphasis)

According to this quote, it is insufficient to know the behavior or properties of the parts of a system "in isolation" because, in order to understand the behavior of a system, we also need to know how these parts are integrated into the system as a whole.[87] This integration is also characterized as a "synthesis step" (Loehle 1988, 100) or as "put[ting] them [i.e. the parts] together again" (Wilson 1988, 270). Keller and Hunter illustrate this point by a picture from Alice's Adventures in Wonderland:

> For 50 years, biologists have focused on reducing life to its constituent parts, first focusing on the cell, then working their way down to the genome itself. However, […] [they] took Humpty Dumpty apart but left the challenge of *putting him back together again*. (Keller 2005, 4; Hunter 2003, 20; my emphasis)

Keller and Hunter state that it is not enough to dissect Humpty Dumpty into its parts and to study these parts. In addition, biologists must "put him together again". Bechtel and Richardson have coined the term '*recomposition*' for this task of putting together, synthesizing, or integrating the parts into a complete, functioning system again (2010, xxxvii-xl). To recompose a biological system implies to show how the parts work together to bring about the behaviors that the system characteristically displays. Bechtel and Richardson describe this as disclosing the "orchestrated functioning" (2010, xxxix) of a system (or mechanism) as a whole. In the biological literature the phrase 'putting the parts together again' is understood as follows:

> Advances in the fields of cybernetics and biology led to the proposition of new interpretative models that were better suited to identifying and describing phenomena that could no longer be seen as abstractly isolated entities divisible into parts…, but needed to be studied in terms of the *dynamic interactions* of their parts. The word system means 'putting together'. Systemic understanding literally means putting things in a context and *establishing the nature of their relationships*... (Grizzi/Chiriva-Internati 2005, 28; my emphasis)

This quotation exemplifies the major criticism biologists pass on reductionism: studying the behavior of a system in a reductionist manner, that is, by investigating the system's parts in isolation, is defective because it does not throw light on how the individual constituents are organized and how they "*dynamically interact*"[88] (Kitano 2002, 1662; my emphasis; see also

---

[87] "[S]hortfalls in reductionism are increasingly apparent. […] So perhaps there is something to be gained from supplementing the predominately reductionist approach with an *integrative agenda*." (Gallagher/Appenzeller 1999, 79; my emphasis; see also Bateson 2005, 35)

[88] The add-on 'dynamically' is superfluous since any interaction is temporally extended, involves changes and, as such, is dynamic and not static.

Chong/Ray 2002; O’Malley/Dupré 2005, 1270; Sorger 2005, 9) in order to produce the behavior of the system as a whole. The following quotation illustrates this major criticism:

> Identifying all the genes and proteins in an organism is like listing all the parts of an airplane. While such a list provides a catalog of the individual parts, by itself it is not sufficient to understand the complexity underlying the engineered object. We need to know *how these parts are assembled* to form the structure of the airplane. This is analogous to drawing an exhaustive diagram of gene-regulatory networks and their biochemical interactions. Such diagrams provide limited knowledge of how changes to one part of a system may affect other parts, but to understand how a particular system functions, we must examine how the individual components *dynamically interact* during operation. (Kitano 2002, 1662; my emphasis)

In Kitano’s view, understanding a biological system (like a gene-regulatory network) resembles understanding an airplane. A list of all parts of the airplane is just as little sufficient for understanding how a biological system works than it is for understanding the functioning of an airplane. In addition, we need to know how these parts are organized and interact with each other.

All in all, this overview about contemporary discussions about reductionism in biology reveals that reductionistic research is associated with two features: first, it counts as reductionistic not only to decompose a system into parts, but also to study the parts in isolation; second, applying this reductionistic strategy results in the failure to capture how the parts are integrated into a system as a whole, that is, how they are organized and interact with each other.

4.1.2 Specifying What Parts in Isolation Are

On basis of this information I can now further specify what the phrase ‘parts in isolation’ means.

Studying Parts in Different Contexts Than *in situ*

One option is that the phrase ‘studying parts in isolation’ implies studying each part of a system *completely on its own*, that is, completely isolated from all other entities. According to this reading, the behavior of a part of a system is explored under different conditions, but not in relation to other entities. This raises the problem that, by investigating the parts in complete isolation, we cannot get information (or, at least, we can get only very limited information) about the relational properties of the parts and how they typically interact with other entities of certain kinds. Hence, if we understood the notion of parts in isolation in this way the shortcomings of the reductionistic research strategy would become intelligible. However, what strongly speaks against this reading is that the procedure of studying an object by isolating it completely from other objects is a research strategy that is not common in biological practice. Instead, in actual experimental settings the different conditions under which the parts of a system are studied usually include other entities.

Therefore, I favor the alternative interpretation, namely to understand the term 'parts in isolation' as 'parts isolated from their original system'. This means that the parts of a system are taken out of their original context and are examined in a different context than *in situ* (for instance, under laboratory conditions, i.e., *in vitro*). These different contexts commonly include other entities, which is why the parts are not studied in complete isolation. The different contexts may even contain other parts of the system, but they need to be different from the context that surrounds the part when it is integrated with the other parts to the system as a whole (i.e. they need to be different from the context *in situ*). The questionable assumption is that the results that one gets from applying this reductive method (i.e. from studying the parts in isolation from their original system) are still valid when the parts are put together again.[89] This assumption turns out to be particularly problematic in the case of so called integrated systems (see Section 1.2.4), whose parts are organized in such a complex way that the properties of the parts and their interactions are co-determined by the systemic organization (more on this in the next section).

So far I focused on the reductionistic strategy to study parts in isolation. Let us now turn to reductive explanations. My central thesis in this section is that to reductively explain the behavior of a system means not only to refer exclusively to lower-level factors (Section 3) and to focus on factors that are internal to the system, that is, on the genuine parts of the system (Section 4). In addition, reductive explanations appeal only to the system's parts in isolation. But how are we to understand this thesis? What does it mean that an explanation cites only parts-in-isolation? Does it imply that a reductive explanation refers only to intrinsic properties of parts, and that the organization of the parts (which can be reframed as relational properties of the parts) and their interactions are ignored altogether?

Referring Only to Some Kinds of Relational Properties and Interactions

I have already touched upon a related issue in Section 2.2.2. There the question was whether the lower-level character of reductive explanations precludes that functional properties of the parts and their organization are represented in the explanation, too. I concluded that the description of functional properties of parts does not violate the lower-level character of explanations (it may, however, violate the internal character). Even the spatial and temporal organization of the parts of a system can be included in a lower-level explanation if these organizational features can be characterized as relational properties of lower-level entities (and not as systemic properties). The question that is at issue in this section is similar to, but distinct from the one addressed in Section 2.2.2. In this section I wonder whether the third feature (and not the first feature, i.e. the lower-level character) of reductive explanations, which I call parts-in-isolation character, is consistent with the fact

---

[89] Wimsatt describes this assumption as the reductionistic strategy to assume that "the results of studies done with parts studied under different… conditions are *context-independent*, and thus still valid when put together to give an explanation of the behavior of the whole". He refers to this heuristic as "Articulation-of-Parts (AP) coherence" (2006a, 470).

that reductive explanations refer not only to intrinsic properties of parts, but also to the relational or functional properties of parts, to how the parts of a system are organized, and to the interactions between the parts.

A philosopher of biology who seems to give a negative answer to this question is Sarkar (his position has been introduced and critically discussed in Chapter III, Section 2.2). In the introduction to his volume on "Molecular Models of Life" (2005) Sarkar amplifies the idea of what can legitimately be invoked in a reductive explanation (more precisely, in the kind of reductive explanation he calls "strong reduction"; 1998, 45). He states that the reductionist has a "more restricted repertoire" (2005, 9) to hand than the antireductionist. More precisely, he claims that in reductive explanations

> the properties invoked in explaining some feature of a whole must be the properties of the parts alone, each *definable without reference to some other part*. (2005, 9; my emphasis)

This quote suggests that reductive explanations refer only to *intrinsic properties* of the parts of a system since intrinsic properties are the only properties that a part has on its own, independent from the other parts of the system.[90] This implies that reductive explanations refer neither to the organization of the system's parts, nor to their interactions because these are relational properties of the parts.[91] This extreme view of reductive explanation and of reductionism in general can only rarely be found in the biological literature. However, a few biologists endorse such a view. For instance, in a paper that applies central insights from systems biology to clinical medicine you can find the following illustration:

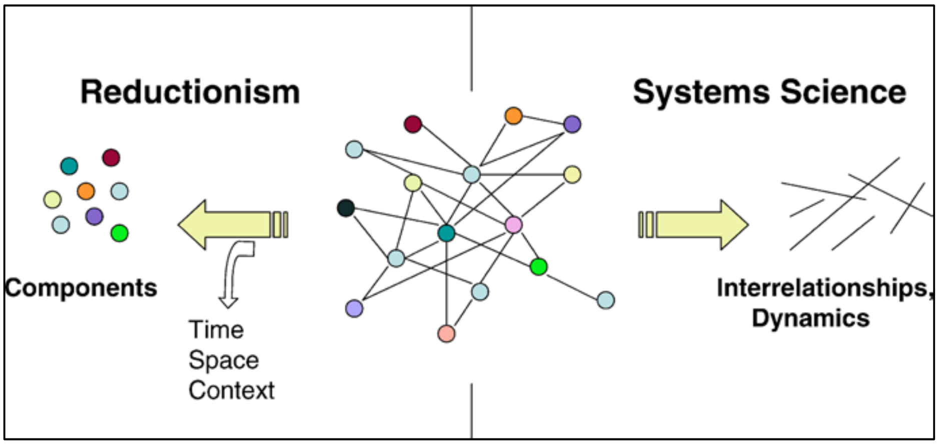

Figure V.21: Reductionism vs. Systems Science in Medicine (Ahn, et al. 2006b, 1)

In accordance with this figure the authors claim that "the act of reduction… disregards component-component interactions and the dynamics that result from them" (Ahn, et al.

---

[90] By contrast, whether an object exhibits a relational (or extrinsic) property depends on the object's relationships to other things.

[91] To be exact, interactions are not relational properties themselves, but rather "occasions on which a change in a [frequently relational] property of one part brings about a change in a [frequently relational] property of another part" (Glennan 2002, 344).

2006b, 1). That is, to investigate and explain the behavior of a system in a reductive way is identified with taking into account only the components of the system with their intrinsic properties and ignoring the relations and interactions between the components.

My impression is that, first, even if the above quote suggests this, Sarkar's other theses clearly show that he does not hold the view that reductive explanations invoke *no* relational properties of and interactions between parts, and that second, independent from what Sarkar claims, it is implausible to restrict the concept of reductive explanation (and of reductionism in general) in such a way that it applies only to explanations, which explain the behavior of a system by referring only to the intrinsic properties of the system's parts.

Let me start with the first point. In Chapter III, Section 2.2 I have introduced and critically discussed Sarkar's account of reductive explanation in genetics. Against this background the above citation seems to be by far too restricted. Let me illustrate this by an example. Sarkar claims that even the phenomenon of cooperativity of hemoglobin, which traditionally has been regarded as a cornerstone of antireductionism (2005, 7f; van Regenmortel/Hull 2002, 254) can be explained reductively (2008, 429; 2005, 7f, 10f, 110). However, the explanation of cooperativity is far from appealing only to intrinsic properties of parts. Consider this example in more detail.

<u>Example: Cooperative Binding of Hemoglobin (Physiology, Molecular Biology)</u>

In order to control a cell's metabolic pathways it is necessary to control the activity of enzymes. One way to do this is allosteric regulation, which subsumes all cases in which a protein's function at one site is affected by the binding of a regulatory molecule (either an inhibitor or an activator) to a separate site. A subtype of allosteric regulation is called cooperativity. In these cases a substrate molecule binds to one site in an enzyme consisting of several subunits and triggers a shape change in all the subunits, thereby increasing catalytic activity at the other active sites. Cooperativity amplifies the response of enzymes to substrates: one substrate molecule primes an enzyme to act on additional substrate molecules more readily.

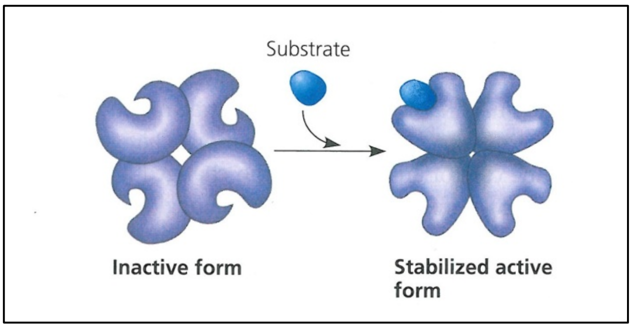

Figure V.22: Cooperativity (Reece, et al. 2011, 204)

The binding of hemoglobin to oxygen in vertebrates is a much discussed example of the phenomenon of cooperativity (even though hemoglobin is not an enzyme). Hemoglobin consists of four subunits (two α-subunits and two β-subunits), each of which has an oxygen-binding site. The binding of an oxygen molecule to one binding site increases the affinity for oxygen of the remaining binding sites. This promotes the binding of oxygen when more oxygen is present (e.g. in the lungs or gills) and it promotes the release of oxygen where it is most needed.

The phenomenon to be explained is the behavior of hemoglobin to raise its affinity to oxygen after having bound an oxygen molecule. This phenomenon is also expressed by the sigmoid binding curve of hemoglobin to oxygen. It is explained by reference to the parts of hemoglobin, namely its subunits (respectively, certain parts of these subunits like heme groups), and how the binding of an oxygen molecule (which is an external factor) affects certain interactions between the subunits/the subunits' parts that brings about a shape change and thereby increases the affinity of hemoglobin to oxygen. In my view, this explanation invokes much more than the intrinsic properties of the parts of hemoglobin. Rather, it appeals to relational properties of the parts as well (e.g. the property of an iron atom of being bound to oxygen or the property of a protein helix of being located close to the iron atom). Many of these relational properties can be characterized as organizational features. Thus, the explanation specifies also how certain parts of the hemoglobin are spatially organized. Moreover, the explanation specifies how different parts of hemoglobin interact with each other and how these interactions are temporally organized. In sum, Sarkar's characterization of the explanation of cooperativity as being reductive as well as his overall position provide considerable doubt that he really wants to claim that reductive explanations cite only intrinsic properties of parts (even though the above citation strongly suggests this).

Second, regardless of what Sarkar claims, I think there are strong reasons why such a restricted notion of reductive explanation is implausible. If one regards only those explanations as reductive explanations which refer exclusively to intrinsic properties of the parts of a system (or of lower-level factors), the result is that there are almost no reductive explanations developed in the biological science. In other words, hardly any adequate explanation of the behavior of a biological system excludes all relational properties of and interactions between the system's parts. Accordingly, to understand the concept of a reductive explanation in this way amounts to building up a straw man that is rarely realized in biological practice, but that is easy to attack.

But note, I am not claiming that reductive explanations appeal to *all* kinds of organization and interactions between the parts of a system. This would be a too broad notion of reductive explanation, which would fail to account for the shortcomings of reductionistic investigative and explanatory strategies that biologists frequently point out. Rather than allowing only intrinsic properties of parts or allowing all kinds of relational properties of and interactions between parts I think we need an intermediate position. It can be obtained by further pursuing the way I have already outlined. The phrase 'parts in isolation' should be specified as 'parts isolated from the original system".

That is, my central thesis is that to explain the behavior of a system by appealing only to the system's parts in isolation means to refer only to those relational properties of the parts and to those kinds of interactions that can be discovered by investigating the parts in other contexts than *in situ*. In other words, reductive explanations include only information that is discovered by taking the parts out of their original system and by exploring the parts in different contexts (e.g. in vitro).

4.2 Treating Systems as Nearly Decomposable

In the biological literature the concept of reductionism is not only associated with studying and explaining the behavior of a system by appealing only to the system's parts in isolation. In addition, two other specifications can be found: on the one hand, reductionism is treated as the view that biological systems are *aggregative systems*; on the other hand, a reductionist is said to take the organization of and interactions between the parts of a system into account, but to represent them as neat, *linear causal chains*. In what follows I will examine these assumptions in more detail and relate them to my thesis about parts in isolation (previous section) and to what I have said about different kinds of decomposability (Section 1.2.4). I will argue that the first assumption (i.e. that a reductionist holds that biological systems are aggregative system) yields not a convincing understanding of reductionism and reductive explanation, whereas the second does. Finally, I will integrate these results with the ones of Section 4.1 in order to identify a third major characteristic of reductive explanation in biology.

Some biologists regard reductionism as the view that biological systems (including organisms) are no more than "the sum of their parts" (Strange 2005, 968), that each of them is "just an assembly of genes and proteins" (Kitano 2002, 1662), and that "the components involved can be added or subtracted from the system" (Bizzarri, et al. 2008, 181). These citations show that some biologists associate reductionism with the assumption that biological systems are *aggregative systems* whose behaviors remain invariant to certain modifications of the system's parts – for instance, to substituting some parts, to reducing the number of parts, and to rearranging them (see Wimsatt 1986, 260-269 and Section 1.2.4). Nagel has termed this the "additive point of view" (1952, 26). This view is problematic since in reality only very few biological systems satisfy the conditions for aggregativity. Thus, to treat biological systems as aggregative systems even if they are not aggregative results in serious errors. If reductionism were to understand in this way the application of reductive research strategies would almost always be inappropriate and reductive explanations of almost all biological phenomena would be inadequate. The conclusion I draw from this is that such an understanding of reductionism is deeply flawed because it amounts to constructing a straw man. And I am not alone with this opinion. Other biologists explicitly deny that reductionism can be identified with the view that biological systems are aggregative systems. For instance, Fincham claims that "[m]olecular biologists in particular tend to be accused of it [reductionism], but they do not hold the *naive view* that complex structures and processes are just sums of their parts" (2000, 343; my emphasis). Similarly, Wilson states that "[d]espite the frequent holistic statement that 'the whole is more than the sum of its parts', reductionists do *not* expect interactions among parts to be *additive*" (1988, 270; my emphasis).[92]

Things stand differently with respect to the second assumption. In the biological literature one finds the view that a reductionist is someone who takes into account the relations and interactions between the parts of a biological system, but who underestimates their complexity because he "analyses complex network interactions in terms of simple causal chains and mechanistic models" (van Regenmortel 2004b, 1018).[93] In other words, a reductionist is said to treat the interactions between the system's parts (i.e. the dynamics of a system) as "linear" (Bizzarri, et al. 2008, 181, 184). Soto and Sonnenschein put this a bit polemically when they state that "those that assume a reductionist stance hope that eventually a *neat, linear causal chain* will be identified" (2006, 372; my emphasis).

In order to understand what this means recall what I have said about different kinds of decomposability in Section 1.2.4. There I pointed out that four kinds of biological systems can be distinguished: aggregative systems that are simply decomposable, component systems that are "nearly decomposable" (Simon 1962, 473), integrated systems that are only minimally decomposable, and non-decomposable systems (see Bechtel/Richardson 2010). What is of particular importance in this context is the difference between component systems and integrated systems. In component systems the properties of the parts are "intrinsically determined" (Bechtel/Richardson 2010, 26). That is, the causal interactions within the parts are more important for determining the part's properties than the interactions that occur between parts. I have already noted that this does *not* imply that the parts of component systems exhibit no relational properties and that they do not interact with each other. Rather, the organization and interactions that do exist are relatively *simple* (this is why Wimsatt calls systems of this kind "interactionally simple"; 1974, 72; 2007, 184). In particular, the contributions that the parts make to the overall behavior of the system are said to be *sequentially* and *linearly*. As a result, organization and interactions provide only secondary constraints on the functioning of constituents (Bechtel/Richardson 2010, 26, 171). By contrast, in integrated systems the organization of the parts and the interactions between them are so complex that they are significantly involved in determining the

---

[92] Philosophers, too, agree with me on this point. For instance, Wimsatt and Sarkar write that in most biological explanations "much of the explanatory weight is borne by the *organization* of… parts into… the higher-level system". Hence, they emphasize that "it is not appropriate for interlevel reduction to be tarred with the ontologically corrosive reputation of aggregativity." (2006, 702)

[93] This underestimation of complexity is a part of what Dennett calls "*greedy reductionism*" (1995, 82).

properties of the parts. This makes the behavior of the parts *mutually interdependent* and restricts their autonomy (Bechtel/Richardson 2010, 170; see also Nagel 1952, 27-29). Accordingly, the components of integrated systems have no "independent, isolable function" (2010, 31), which is why they cannot be studied in isolation.

The difference between component and integrated systems provides us with a new way to think about the phrase 'parts in isolation' that also clarifies how this phrase is related to the assumption that a reductionist regards biological systems as if their parts exhibited only simple, linear interactions. My proposal is to link reductionism to the view that biological systems can be investigated and explained as if they were component systems, that is, as if they were nearly decomposable. In other words, to investigate the behavior of a system in a reductive manner means to treat systems as if they were composed of parts, whose properties are mainly unaffected by their context and which, thus, can be taken out of the system and studied in isolation (i.e. in other contexts than *in situ*). Furthermore, to reductively explore a system implies to assume that the organization of the parts of the system and the interactions between them are relatively simple, for instance, that the interactions are sequentially and linearly. In short, the reductionist methodology can be characterized as treating a system as if it were nearly decomposable.

On the basis of this characterization it becomes also clear what the *shortcomings* of employing *reductionistic strategies* are. If the system of concern is an integrated rather than a component system, it may result in serious errors to study the parts of the system in isolation and to assume that the parts make an independent, linear contribution to the behavior of the system. The limitations of the reductionistic methodology are primarily due to the fact that in integrated systems the properties of the parts are co-determined by the systemic organization. Consequently, the parts often exhibit different properties when they are integrated into the system (i.e. *in situ*) and when they are taken out of their original system and studied in different contexts. This fact is also stressed in the biological literature on reductionism. For instance, in their introduction to a special issue on systems biology as "Whole-istic Biology" Chong and Ray argue that "'systems' of various orders [are] not understandable by investigation of their respective parts in isolation" because of "the difference of behavior of parts when isolated or in higher configuration" (2002, 1661; see also Keller 2005, 9).

This way of characterizing the reductionistic methodology points to a third feature that is characteristic for reductive explanations in the biological science. Reductive explanations do not only refer exclusively to lower-level factors (Section 2) and focus on those factors that are internal to the system in question (Section 3), they also appeal only to the system's parts in isolation. The latter characteristic means that reductive explanations describe only those relational properties of and interactions between the genuine parts of a system that can be discovered by studying the parts in other contexts than *in situ*. This implies treating the system of interest as a nearly decomposable system, that is, as a system whose genuine parts are organized in a simple way, interact with each other in a sequential and linear fashion, and exhibit properties that are largely unaffected by their context (and can, thus, be studied in isolation).

5 Part-Whole, Mechanistic, and Reductive Explanation

The core of my account of explanatory reduction can be summarized as follows. Reductive explanations exhibit two necessary features. They explain the behavior of a biological system by referring only to

(1) factors that are located on a *lower level* of organization than the system, and to
(2) *parts in isolation* (i.e. to those relational properties of and interactions between parts that can be studied in other contexts than *in situ*).

Moreover, many (but not all) reductive explanations exhibit also a third characteristic. They explain the behavior of a biological system by focusing on

(3) factors that are *internal* to the system (i.e. that are genuine parts of the system).

There are many reasons in favor of the view that, taken together, these three conditions are sufficient for a biological explanation to be reductive.[94]

Since reductive explanation is said to be closely connected to part-whole explanation and to mechanistic explanation the question arises how exactly the relation between these three kinds of explanation can be specified. I have already touched on this issue in the previous sections. However, in this section I will explore this question in detail.

Consider part-whole explanation first. Part-whole explanations are paradigmatic cases of reductive explanations, which is why some philosophers identify reductive explanations with part-whole explanations (e.g. Hüttemann/Love 2011). I have already argued that this is a mistake. Even though the set of part-whole explanations and the set of reductive explanations overlap to a great extent, it is important to note that they are not congruent. In Section 3.1 pointed out that all explanations, which explain the behavior of a system by appealing exclusively to factors that are internal to the system (i.e. that are genuine parts of the system) are part-whole explanations. However, I also argued that for an explanation to be reductive it need not invoke only the genuine parts of a system. It may also refer to environmental factors (also called contextual or external factors) provided that these environmental factors are located on a lower level than the system whose behavior is to be explained (see Section 2.2.1). In short, the internal character is *not* a *necessary condition* for the reductivity of an explanation, whereas the lower-level character is. The lower-level and the internal character can fall apart if we accept a global notion of level, on

---

[94] However, I do not insist on this thesis. Biological practice exhibits a sometimes surprising diversity. Perhaps in some biological field there exists a kind of reductive explanation I have overlooked, but which is crucial to that field. Future work on this issue will show whether my list is complete or whether it must be supplemented by other characteristics.

basis of which level claims are not restricted to the parts of the system of interest, but can also be made about contextual factors (see Section 3.2). Reductive explanations that exhibit a lower-level character, but not an internal character, are cases of reductive explanations that are *no part-whole explanations*. This is the respect in which the category of reductive explanations is broader than the category of part-whole explanations.

In turn, the category of part-whole explanations is also broader than the category of reductive explanations because not all part-whole explanations satisfy the third condition of reductivity I discussed in Section 4. That is, not all part-whole explanations appeal only to those relational properties of and interactions between the parts that can be discovered by studying the parts in isolation (in other words, by treating the system as a nearly decomposable system). This is exactly what biologists who call for a move "beyond reductionism" (Gallagher/Appenzeller 1999, 79) are driving at: the behavior especially of complex systems cannot be adequately explained by taking into account only the parts in isolation. Rather, the dynamics of the system and the way the parts are integrated into the system as a whole must be considered. Explanations that satisfy these demands may be part-whole explanations, but not reductive explanations. However, if the system is entangled in an inseparably complex way with its context it might be that in order to develop an adequate explanation of its behavior certain contextual factors must take center stage in the explanation. As a result, the explanation even fails to be a part-whole explanation (i.e. to possess an internal character).[95]

The following table sums up the relation between reductive explanation and part-whole explanation:

	reductive explanation	**part-whole explanation**
lower-level character	necessary	not necessary
internal character	not necessary	necessary
parts-in-isolation character	necessary	not necessary

Table V.23: Comparing Reductive and Part-whole Explanations

Let us turn to the relation between reductive explanation and mechanistic explanation. In discussions about reductionism in the biological sciences these two concepts are almost identified. For instance, in cancer research the "shortcomings of the reductionistic… thought" are mentioned in the same breath with the "shortcomings of the… mechanistic thought" (Bizzarri, et al. 2008, 180). Likewise, Sorger states:

> Molecular biology emphasized reductionist approaches to carefully delineated problems on the premise that important insights derive from deep mechanistic understanding. (2005, 9)

---

[95] Alternatively, it might be that the complex entanglement of system and environment results in a failure of explanation altogether because the system cannot be separated from its environment and, thus, the phenomenon to be explained cannot be identified (see Section 1.1).

And in the following quote reductionism is explicitly equated with "the mechanistic movement":

> Systemism was born in the first half of the twentieth century as a reaction to the previous mechanistic movement (also known as reductionism). (Grizzi/Chiriva-Internati 2005, 28)[96]

However, one should note that the concept of mechanism that these biologists have in mind is not the one that is discussed in the context of the "New Mechanistic Philosophy" (Skipper/Millstein 2005, 327) in contemporary philosophy of science. Rather, their understanding of 'mechanism' traces back to the Mechanistic Philosophy of the late 16th and 17th century and to the subsequent dispute about mechanicism and vitalism in the 18th and 19th century. This also makes clear why they connect the notion of mechanism so closely to the notion of reduction. In the Mechanistic Philosophy of the late 16th and 17th century René Descartes, Robert Boyle, and others argued that all natural phenomena can be made intelligible by mechanistically explaining them in terms in terms of material particles, their few, simple properties (like size, shape, and motion), and the interactions between them.[97] Although, at that time, nobody used the term 'reduction' or 'reductive explanation' the mechanistic explanations Descartes and Boyle favored are clear cases of reductive explanations, more precisely of fundamental-level explanations.[98]

In contemporary philosophy of science the Mechanistic Philosophy experiences its second spring. However, since the 16th and 17 the century the concept of mechanism and of mechanistic explanation has changed a lot. Nowadays, mechanistic explanations are not only restricted to explanations that appeal to the size, shape, and motion of material particles. Rather, mechanistic explanations are allowed to invoke many different kinds of entities and their properties as well as many different forms of activities/interactions. On which level a mechanistic explanation "bottom[s] out", that is, which components are "accepted as relatively fundamental or taken to be unproblematic" (Machamer, et al. 2000, 13), is relative to the explanatory interests and purposes of a given scientist, research group, or field. This, however, does not imply that mechanistic explanations no longer possess a lower-level character. The opposite is true. The New Mechanists still treat mechanistic explanations as being *lower-level explanations* (or, as some of them term it, "interlevel explanation"; Bechtel 2008, 148-155) that span two levels, namely the level of the mechanism as a whole that shows the behavior to be explained and the level of the components of this mechanism. The notion of mechanistic explanation is just not confined to the class of fundamental-level

---

[96] See also van Regenmortel 2004b, 1018; O'Malley/Dupré 2005, 1270; and Bizzarri, et al. 2008, 184.

[97] See for instance Boyle, "The Excellency of Theology" (1966), 69-71, 77f and Descartes, "Principles", Part III, Section 46. This traditional idea can still be found in the statements of some biologists. For instance: "Reductionism seeks to explain the wide variety of natural phenomena on the basis of the behavior of a *limited number of simple constituents*" (Grizzi/Chiriva-Internati 2006, 5; my emphasis).

[98] Note that the term 'fundamental level' here means not the lowest biological level of molecules, but rather the lowest physical level that was assumed at that time (i.e. the level of material particles and of their properties size, shape, and motion).

explanations anymore (with the fundamental level being the level of material particles with their properties size, shape, and motion).

Before I consider the other two characteristics of reductive explanation, let me add a general remark. In the literature on mechanisms an *ambivalent stance on reduction* can be observed. On the one hand, some mechanists either explicitly distance themselves from reductionistic views about scientific practice (Craver 2005; 2007a, 196-171) or avoid using the concept of reduction at all (Bechtel/Richardson 2010, xiii, xvii, xxxvii-xl). On the other hand, other mechanists make such ambitious claims like:

> Thinking in terms of mechanisms provides a new framework for addressing… [the] traditional issue… [of] reduction. (Machamer, et al. 2000, 1)

This seeming contradiction can be dissolved by recognizing that those mechanists who oppose reduction or avoid talking about it criticize a specific kind of reduction, namely theory reduction (see Chapter II, Section 3.1 and 4). They do not oppose the idea of reduction in general.[99]

Hence, the New Mechanists appear to maintain the old alliance between mechanisms and reduction. But how far does this alliance go? How are we to interpret Wimsatt's thesis that "explaining types of phenomena by discovering mechanisms… is seen by them [i.e. by biologists] as reduction, or as integrally tied to it" (1976a, 671)? In other words, is the set of all mechanistic explanations really coincident with the set of all reductive explanations? In what follows I successively examine whether mechanistic explanations exhibit the three characteristics of reductive explanations. First, I have already argued that mechanistic explanations satisfy the first condition of reductivity, that is, since they refer only to the components of a mechanism they appeal only to lower-level factors. However, just as in the case of part-whole explanations it must be noticed that the category of mechanistic explanations is narrower than the category of lower-level explanations. The latter also includes explanations that also invoke lower-level factors that are not components, but belong to the context of the mechanism in question.

Second, consider now the second feature of reductive explanations, their internal character. In Section 1.2.1 I have pointed out that the notion of a component of a mechanism needs to be distinguished from the notion of a genuine part of a system. All entities and interactions that first, are located inside the spatial boundary of the system S, and that second, contribute to bringing about at least one of behaviors that S typically displays are genuine parts of S. By contrast, only those entities and interactions (or activities) are components of a mechanism that work together to bring about a *specific* behavior of the corresponding system. Accordingly, mechanistic explanations appeal only to a subset of those factors that are internal to (i.e. genuine parts of) the corresponding system (but to all

---

[99] This is apparent, for instance, in Machamer, et al. 2000, 23; Bechtel/Richardson 2010, xxxvii; and Bechtel/Hamilton 2007, 405.

components of the mechanism). But this does not violate their internal character. Another peculiarity of mechanistic explanations is that they might appeal also to environmental factors. This traces back to the somewhat peculiar thesis of the mechanists that “set-up” or “enabling conditions” and even background conditions are regarded as “parts of the mechanism” (Machamer, et al. 2000, 11; see also Section 1.3.3). However, most set-up conditions, like the extracellular neuronal signal that triggers the intracellular process of muscle contraction, are not genuine parts of the system and as such external to it. Accordingly, a mechanistic explanation might also cite contextual factors. But since the included contextual factors are simplified as being input or background conditions, the internal character of mechanistic explanation is preserved. In conclusion, the internal character is a necessary condition for an explanation to be mechanistic, whereas it is only a typical feature that many but not all reductive explanations possess.

Third, in the biological literature the reductive strategy of studying and explaining the behavior of a system by reference to the system’s parts in isolation is characterized as analyzing “complex network interactions in terms of simple causal chains and mechanistic models” (van Regenmortel 2004b, 1018). In accordance with that, some mechanists focus in their work on mechanisms with relatively simple organizations (e.g. Glennan 2002, 2008; Tabery 2004). However, due to their “expansionist tendencies” the majority of mechanists argue that their account equally applies to more complex kinds of systems, like integrated systems, which exhibit complex forms of organization like feedback (see, in particular, Bechtel/Richardson 2010; Bechtel 2001, 2009; Bechtel/Abrahamsen 2011). In other words, they do not restrict the kinds of organization of and interactions between parts that are allowed in mechanistic explanation to those that can be discovered by studying the components of the mechanism in isolation (i.e. in other contexts than *in situ*). Thus, if we accept their account, the result is that the set of mechanistic explanations contains not only reductive explanations. In addition, it encompasses also non-reductive explanations, which possess a lower-level and an internal character, but do not meet the third condition of reductivity, that is, they do not only appeal to parts in isolation.

The following table sums up the relation between reductive explanation and mechanistic explanation:

	reductive explanation	**mechanistic explanation**
lower-level character	necessary	necessary
internal character	not necessary	necessary
parts-in-isolation character	necessary	not necessary

Table V.24: Comparing Reductive and Mechanistic Explanations

Finally, have a look at a thesis that steps out of the line. In his work Bechtel repeatedly argues in favor of the concept of “mechanistic reduction” (2008, 157; Bechtel/

Hamilton 2007, 410) or “mechanistic reductionism” that “neither denies the importance of context or of higher levels of organization nor appeals exclusively to the components of a mechanism in explaining what the mechanism does” (Bechtel 2006, 41). The latter part of this citation is unproblematic. Of course, a mechanistic explanation does not merely describe the components of a mechanism (including their intrinsic properties). Furthermore, any mechanistic explanation refers also to the relations between these components and to the way on which they interact with each other (respectively, which activities they perform). It is the former part of this quote (i.e. the statement that mechanistic explanations appeal to the context and to higher levels) that is problematic. This thesis cannot be dismissed as a marginal note. Rather, it is central to Bechtel’s work. This becomes manifest for example in the title of a recent paper on “Looking down, around, and up: Mechanistic explanation in psychology” (2009). In this and other papers Bechtel claims that in order to develop an adequate mechanistic explanation researchers also must “look up – situate the mechanism in its context, which may be a larger mechanism that modulates its behavior” (2009, 543). In other words, a mechanistic explanation must take into account how the mechanism as a whole interacts with entities in its environment and how the mechanism is embedded into higher-level mechanisms (2006, 41-43; 2008, 155-157). According to Bechtel’s account, mechanistic explanations fail to be reductive with respect to all three characteristics of reductive explanations in biology. He concurs with other mechanists that mechanistic explanations do not refer only to parts (or components), but he adds that mechanistic explanations also do not appeal exclusively to lower-level factors and to internal factors. However, Bechtel is the only one who holds this radical and, as I think, not really convincing view.

6 The Ontic Character of My Account

As the title of my book indicates, I characterize my analysis of reductive explanation, the core of which I have just presented in the preceding sections, as an “ontic” account of explanatory reduction. This gives rise to the question of what it is that makes my account ontic. In Chapter IV, Section 1.3 I made the first steps towards answering this question. There I spelled out which features constitute the ontic character of an account of explanation. I argued that similar reasons can be adduced for why my account of explanatory reduction should be called ontic, too. Let me briefly recall what it is that makes an account of explanation ontic.

I start with a negative thesis. I do not agree with philosophers like Craver or Strevens that what makes an account of explanation ontic is the thesis that explanations are objects or facts that exists in the world itself, independently from scientists seeking, discovering, and representing them. Rather, I think that explanations are epistemic units (e.g. descriptions, representations, models, or explanatory texts). My positive thesis is that characterizing a conception of explanation as “ontic” indicates a certain thesis about what determines whether an explanation is adequate or not (respectively, whether a representation is explanatory or not). According to an ontic account of explanation these determinants of explanatory force are *not logical relations* between explananda-sentences and explanantia-sentences, but rather *(causal) relations that exist in the world*. That is, the ontic character of a conception of explanation can be linked to the thesis that the fact that a description truly represents certain (causal) facts in the world decides on that it is explanatory. In other words, ontic accounts trace the difference between explanations and non-explanations, ultimately, back to differences in the world, and not to logical features of linguistic entities. Forge puts this as follows:

> "the ontic conception… amount[s] to the claim that… what makes an explanation work, is entirely a matter of *what the world is like*" (Salmon 1998, 77; my emphasis).[100]

This understanding of the term 'ontic' can be applied to my analysis of reductive explanation in biology. My account of explanatory reduction is ontic, too, because it traces back the difference between reductive explanations and non-reductive explanations to particular relations that exist in the world, rather than to certain logical relations between explanandum-sentence and explanans-sentence. My central claim is that what determines whether an explanation is reductive or not is whether it truly represents specific relations or facts that exist in the world. Contrary to proponents of a Nagelian model of theory reduction (see Chapter II, Section 3.1 and 4) I regard formal issues (like the deductive relation between explanandum-sentence and explanans-sentence) to be less relevant to reduction. Another way to frame the ontic character of my account is thus to say that my account focuses not on formal, but on what Sarkar calls "substantive issues" (1998, 39) of reduction.

In order to explain what it means that my account traces back the reductivity of an explanation to particular relations that exist in the world, recall the three characteristics of reductive explanations that I presented in this chapter. The core of my account of explanatory reduction in biology is that in reductive explanations the behavior of a system is explained by, first, appealing only to factors that are located on a *lower level* of organization than the system, second, by focusing on factors that are *internal* to the system (i.e. genuine parts of the system), and third, by describing the system's *parts in isolation* (i.e. by describing them as objects, whose properties are not determined by the systemic organization and that interact with each other in a linear, sequential way). I argued that the first and third of these features are necessary conditions, whereas the second is only a typical characteristic that many reductive explanations exhibit. These three major features of reductive explanations demonstrate the ontic character of my account. According to this view, what determines the reductivity of an explanation is whether they truly represent

---

[100] When characterizing what the ontic character of a model of explanation is, other authors highlight more the *absence of deductive arguments*. For instance, Carruthers states that an account of explanation is 'ontic' if "there is no attempt to construct deductive arguments in which the explanandum figures as the conclusion. Rather, explanations proceed by telling us about the causes or constitution of their targets" (2004, 160).

certain relations or facts that exist in the world. I identify three of these relations: first, the relation that one factor is located on a higher, on a lower, or on the same (global) *level of organization* than another factor (1); second, the relation that a factor is internal to a biological system, that is, that the factor is located *inside of the spatial boundary* of a system (2); and third, the fact that a system consists of *parts in isolation*, that is, of objects whose properties are not determined by the systemic organization and that interact with each other in a linear, sequential way (3). If an explanation satisfies the above conditions, but is not true (for instance, because the properties of the genuine parts of the system of concern are affected by the way they are organized and by the interactions with other parts or because the interactions between the parts is not sequential and linear), not the *reductivity* of that explanation is called into question, but its *adequacy*.

It is important to notice that the ontic character of my account does *not* convert my analysis of explanatory reduction into an analysis of *ontological reduction*. In Chapter II, Section 2 I pointed out that ontological reduction is a relation between objects, properties, or processes that exist in the world itself. By contrast, explanatory reduction is a subtype of epistemic reduction. That is, it is a relation that holds between parts of our knowledge about the world (and not between parts of this world itself). In case of reductive explanations these epistemic units between which the relation of reduction holds are the representation of the phenomenon or behavior of a system to be explained (explanandum) and the representation of the explanatory relevant factors (explanans). Even if I my account is "ontic", it remains an account of *epistemic* reduction (more precisely, of explanatory reduction). My central question is what renders explanations reductive, not how the putatively reductive relation of constitution (e.g. between a system and it's organized, interacting parts) can be spelled out. What renders my account of explanatory reduction "ontic" is my claim that the reductive character of an explanation is due to the fact that it truly represents specific relations or facts that exist in the world.[101] However, this does not imply that the world *completely* determines the reductivity of an explanation. As I have argued in Section 1.1.3, pragmatic factors may have an influence on how the line between a system and its environment is drawn and on which entities and interactions are identified as genuine parts of a system. Thus, they will also affect the reductive character of an explanation.

7 Interim Conclusion

In this chapter I developed my own account of explanatory reduction in biology. I proceeded in two major steps. The goal of the *first part* of this chapter (Section 1) was to specify four concepts that are central to my analysis of the features of reductive explanations, namely

---

[101] These relations can but need not be relations of ontological reduction. For instance, relations between levels as well as internal-external relations trace back to constitutional relations between a system and its parts. These may be characterized as instances of ontological reduction (see Chapter II, Section 2.1.1).

the concept of a biological system and of its context or environment (Section 1.1), the concept of decomposing a system into its parts (Section 1.2), and the concept of levels of organization (Section 1.3). These conceptual clarifications provided the basis on which, in the *second part* of this chapter (Section 2 to 6), I could develop the core of my account of explanatory reduction. My main thesis was that reductive explanations in the biological science exhibit three characteristics. They explain the behavior of a biological system, first, by appealing solely to factors that are located on a lower level than the explanandum phenomenon (Section 2), second, by focusing on factors that are internal to the system of interest, that is, that are genuine parts of the system (Section 3), and third, by referring only to the parts of a system in isolation (Section 4). The first and third characteristics are necessary for an explanation to be reductive, whereas the second characteristic is only one that most reductive explanations in biology exhibit. There is much evidence that, taken together, these three conditions are sufficient for the reductivity of an explanation. After this rough overview let us now consider the results of this chapter in more detail.

The goal of the *first part* of this chapter (Section 1) was to specify those concepts that are crucial to specifying the notion of reductive explanation in biology, namely the concept of a biological system (Section 1.1), the concept of a part (Section 1.2), and the concept of a level of organization (Section 1.3).

In *Section 1.1* I argued that biological systems are composed of sets of interacting entities that together form a unified whole. This means that the parts of a biological system satisfy two conditions: on the one hand, they work together to bring about those behaviors that the system characteristically displays (Section 1.1.1), and on the other hand, they are located within a defined zone, which is surrounded by a spatial boundary. I claimed that the spatial boundary that separates a system from its environment ideally is a continuous, constant surface, but that this is not always the case. The boundaries of some systems are subject to continuous change. Others are not continuous surfaces, but rather conglomerates of several disconnected boundaries. I revealed that it sometimes might be difficult to identify the boundary between a system and that biologists do not always agree on how to perform this task. However, I stressed that in most cases this does not imply that biologists *fail* to distinguish system and environment (Section 1.1.2). Finally, I rejected the view that the separation of a system from its context is a highly flexible or even an arbitrary matter. I argued that, instead, the boundaries of many biological systems fall together with certain structures that exist in nature (e.g. with membranes), which is why I characterized them as bona fide boundaries, rather than as fiat boundaries (Section 1.1.3).

In *Section 1.2* I analyzed what it means to decompose a system into its parts. I began with distinguishing three notions of parts, namely the concept of a mereological part, of a genuine part, and of a component. I concluded that with respect to biological systems the notion of a genuine part is the one that is most adequate. The notion of a mereological part is too unrestricted. And the notion of a component overlooks two important facts, namely that the parts of a system are individuated with respect to *all* behaviors that a system typically displays (and not only with respect to one) and that the parts of a system are located inside the spatial boundary of the system (Section 1.2.1). With regard to the ontological nature of parts I argued that it is important not to confine the notion of a genuine part to spatially located objects since systems are composed of temporally extended interactions or activities, too (Section 1.2.2). Subsequently, I specified the notion of decomposition in two ways: on the one hand, I rejected the pluralistic view that different theoretical perspectives or partitioning frames result in different, equally adequate, non-coincident, and (partly) inconsistent decompositions of the same system. I adopted a more optimistic view and stated that the different decompositions of a system frequently are consistent and can be integrated via other compositional relations (Section 1.2.3). On the other hand, I distinguished different kinds of decomposability and corresponding kinds of biological systems: aggregative systems are simply decomposable, component systems are nearly decomposable, and integrated systems are minimally decomposable (Section 1.2.4).

Section 1.3 served to clarify the notion of levels of organization that my analysis of reductive explanation presupposes. The result of my critical examination of different accounts of levels is that, first, levels of science (e.g. of theories) are a deceptive indicator for levels of nature since different theories from different disciplines refer to entities at the same level and since often one theory appeals to entities from different levels (Section 1.3.1). Second, I argued that size can be no more than a rough indicator of levels, too, because size relations are only by-products of compositional relations, which are the proper relations that determine levels (Section 1.3.2). Third, I emphasized, however, that it must be specified *which* part-whole relations give rise to different levels of organization. I showed that mereological-part-system relations are inadequate, just as Craver's component-mechanism relations (Section 1.3.3). I claimed that, instead, genuine-part-system relations seem to be the one that give rise to a convincing notion of levels of nature (Section 1.3.4). Finally, I pointed out that the notion of level that underlies discussions about reductionism and reductive explanation in biology cannot be captured by a concept of locally defined levels (Section 1.3.3). Rather, I argued that my analysis needs an account of global levels. In I assumed that such a global notion of level may be obtained, for example, by bringing together different genuine-part-system distinctions and establishing a hierarchy of different genuine-part-system relations (Section 1.3.4).

In the *second part* of this chapter (Section 2 to 6) I applied the concepts and distinctions that I had introduced in the first part. The central question that I addressed in this second part was: Which characteristics are common to reductive explanations in the biological science and distinguish them from non-reductive explanations? My answer, which constitutes the core of my account of explanatory reduction in biology, can be summarized as follows: I argued that reductive explanations exhibit two necessary features, namely they explain the behavior of a biological system by referring only to

(1) factors that are located on a *lower level* of organization than the system, and to
(2) *parts in isolation* (i.e. to those relational properties of and interactions between parts that can be studied in other contexts than *in situ*).

Many (but not all) reductive explanations also exhibit a third characteristic. They explain the behavior of a biological system by focusing on

(3) factors that are *internal* to the system (i.e. that are genuine parts of the system).

I claimed that there are many reasons in favor of the view that, taken together, these three conditions are sufficient for a biological explanation to be reductive.

After disclosing these three features I concluded by pointing out the similarities and differences between reductive, part-whole, and mechanistic explanation (Section 5) and by elucidating the ontic character of my account of explanatory reduction (Section 6). Let us consider the findings of the second part of this chapter in more detail.

In *Section 2* I presented the first characteristic of reductive explanations, namely their lower-level character. I started with analyzing examples of reductive explanations from the biological discipline that is regarded as the embodiment of the success of reductionism, namely molecular biology (Section 2.1). I specified the lower-level character of reductive explanations in two ways (Section 2.2): on the one hand, I pointed out that reductive explanations are characterized by an unidirectional flow of explanation from the lower to the higher level, which is why the appeal to downward causation renders an explanation non-reductive (Section 2.2.1). On the other hand, I clarified what exactly it means for an explanation to refer solely to lower-level factors. It turned out that three issues are of particular importance in this context: first, I argued that the set of lower-level factors does not only contain the genuine parts of the system under consideration. It comprises also such factors that are external to the system, but are located on the same level(s) as the genuine parts of the system. Second, I claimed that the category of lower-level factors encompasses also functional properties of lower-level objects if the context that needs to be included is spelled out in lower-level terms, too. Third, I stated that the organization of the system's parts counts as a lower-level factor if it can be characterized as a lower-level (relational) property and not as a systemic property (Section 2.2.2). I introduced two significant subtypes of lower-level explanations, namely fundamental-level explanations and single-factor explanations. I characterized fundamental-level explanations as such reductive explanations that refer only to factors that are located on the lowest biological level (i.e. the "fundamental" level), namely the level of molecules (thus, fundamental level explanations are molecular explanations, which include genetic explanations). I argued that, although molecular explanations are an important subtype of reductive explanations, philosophers should refrain from identifying reductive explanations with fundamental-level explanations because this obscures the diversity of reductive explanation in the biological science and converts explanatory reductionism into a straw man (Section 2.3.1). In single-factor explanations the behavior of a system is explained by reference only to a single causal factor.

I revealed that reductive explanations of this kind are inadequate with respect to most biological phenomena, which is why they rarely occur in biological practice (Section 2.3.2).

Section 3 introduced a second feature that is typical (but not necessary) for reductive explanations in the biological sciences, namely that they explain the behavior of a system by focusing on those factors that are internal to this system. I argued that any factor that is internal to a particular system is a genuine part of this system (Section 3.1). The internal character of reductive explanations has been overlooked or intermingled with their lower-level character for a long time. This is why I think it is important to be aware of the interdependencies and the differences between these two features of reductive explanations. With respect to these differences I concluded that the category of lower-level factors is broader than the category of internal factors. This is to say that the set of all part-whole explanations or explanations with an internal character constitutes only a subset of the set of all lower-level explanations (Section 3.2). Finally, I elaborated on the thesis that reductive explanation focus on internal factors. My main thesis was that this implies that either environmental factors are ignored altogether (i.e. not mentioned in the explanation at all) or that they are simplified to a great extent. I argued that the context of a system can be simplified in two major ways: on the one hand, environmental factors can be simplified as background conditions, which are treated as being homogeneous or remaining constant over time. On the other hand, they can be simplified as mere input conditions, which need to be satisfied only at the beginning of the process that brings about the explanandum phenomenon “automatically” and whose specific character and origin is regarded as irrelevant (Section 3.3).

In *Section 4* I presented a third characteristic of reductive explanations in biology, namely that they explain the behavior of a system by appealing exclusively to the system’s parts in isolation. I argued that philosophers have not paid much attention to this necessary feature of reductive explanations yet, but that the current vivid discussions in the biological literature about the limits of reductionism demonstrate that they should (Section 4.1.1). Subsequently, I considered the question of how exactly the phrase ‘parts in isolation’ is to be understood. I identified two options, of which I chose the second. First, to study parts in isolation can mean to study each part of a system completely on its own. I argued that this interpretation fails to correspond to actual experimental practice in biology. The second interpretation, which I favored, is to understand the phrase ‘study parts in isolation’ as referring to investigating the parts in isolation from their original system (whose behavior is to be explained), that is, in other contexts than *in situ*. I further elaborated this thesis by pointing out that it does not mean that in reductive explanations only intrinsic properties are described. I argued that such a restricted view of reductive explanation would amount to construing a straw man. I claimed, however, that reductive explanations do not refer to *all* relational properties of parts and to the interactions between them, but only to those that can be discovered by investigating the parts in other contexts than *in situ* (Section 4.1.2).

Finally, I examined the relation between this claim and two other claims, namely that the reductionistic methodology amounts to treating biological systems, first, as aggregative systems or, second, as nearly decomposable systems. It turned out that the first view results in a too restricted und untenable notion of reductionism, but that the second view yields valuable insights into the third feature of reductive explanations. I concluded that appealing to parts in isolation can be rephrased as describing the genuine parts of a system as being organized in a simple way, as interacting with each other in a sequential and linear fashion, and as exhibiting properties that are largely unaffected by their context (Section 4.2).

The goal of *Section 5* was to clarify the similarities and differences between three kinds of explanation, namely between reductive, part-whole, and mechanistic explanation. The main results are summarized in the following table:

	reductive explanation	**part-whole explanation**	**mechanistic explanation**
lower-level character	necessary	not necessary	necessary
internal character	not necessary	necessary	necessary
parts-in-isolation character	necessary	not necessary	not necessary

Table V.25: Comparing Reductive, Part-whole, and Mechanistic Explanations

Finally, in *Section 6* I explained why I call my account of explanatory reduction "ontic". I argued that this denotes that I trace back the difference between reductive explanations and non-reductive explanations to particular relations that exist in the world, rather than to certain logical relations between explanandum-sentence and explanans-sentence. In other words, according to my account of explanatory reduction what determines whether an explanation is reductive or not is whether or not it truly represents specific relations or facts that exist in the world, namely whether the explanatorily relevant factors are described as being located on a lower level than the system whose behavior is to be explained, whether they are represented as being internal to the system, and whether they are described as being parts in isolation.

Conclusion

In this brief concluding chapter I summarize what I think are the take-home messages of this book. Furthermore, I sketch some prospects for future philosophical work on the topic of epistemic reduction in the biological sciences.

The central aim of this book was to provide an understanding of an important element of contemporary biological research practice, namely of *explanatory reduction*, respectively of reductive explanation. My main question was: what are the characteristics of reductive explanations in biology that determine their reductivity? In other words, which features are common to all (or to most) reductive explanations in the biological sciences and allow one to clearly distinguish reductive from non-reductive explanations? In Chapter V I answered this question by developing my own account of explanatory reduction in biology. Accordingly, this is also the place in my book where the most notable, novel insights can be found. However, the other four chapters do not only contain preliminary remarks, which can be ignored here. By contrast, they yielded remarkable findings, too, the most important one of which shall be repeated here. In what follows I will at first consider these findings (Chapter I to IV) before I then proceed to the main results of my analysis of the reductive character of biological explanations (Chapter V).

The primary goal of my meta-philosophical analysis in *Chapter I* was to disclose the aim, methodology, and criteria of adequacy of my own analysis of reductive explanation in biology. That way, it secondarily contributed to clarifying how philosophy of science in general can or should be pursued. One important result was that different kinds of projects can be distinguished in philosophy of science. Recall the figure that illustrates these differences:

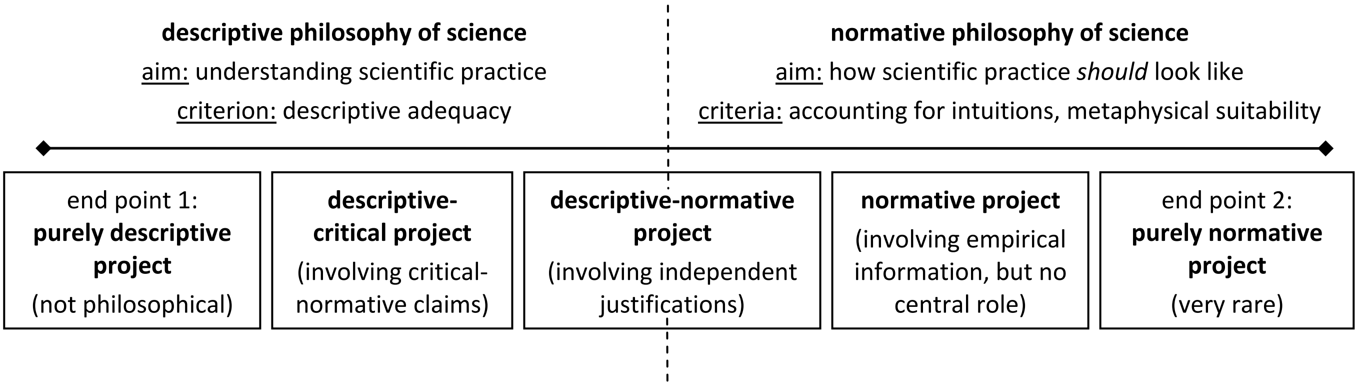

Figure I.2: A Continuum of Different Kinds of Projects in Philosophy of Science

The most important kinds of projects are the three in the middle of this continuum. I characterized my own project of developing an account of explanatory reduction in biology as a project of the descriptive-critical kind (Section 3 and 5.4). With respect to normative projects I argued that they are distinct from but equally legitimate as descriptive-critical projects (Section 2). I stated that things stand differently with regard to projects that claim to take biological practice seriously but that, nevertheless, contain independent justifications of epistemic norms. In my view, these descriptive-normative projects encounter serious difficulties (Section 5). This classification also allowed me to distinguish three different dimensions of normativity (Sections 2.2 and 5). I argued that philosophical accounts can be normative in three different respects: because they aim at

(1) developing an account of how science *should* ideally be carried out or how a scientific concept *should* be understood, without taking into account how science actually works (normative projects), or
(2) understanding actual scientific practice, but critically reconstruct the empirical data that scientific practice provides them with (this involves normative claims about which data *should* be included into the account) (descriptive-critical projects), or
(3) understanding actual scientific practice (in particular, the in fact accepted epistemic norms), but also deliver an independent justification of those epistemic norms that *should* be accepted in science (descriptive-normative projects).

I characterized my own account as being normative in the second, critical sense. What is more, I argued that philosophical projects in general cannot be purely descriptive, but *must* involve a critical-normative stance on the empirical data that scientific practice provides them with (Section 3). I claimed that this critical stance amounts to, first, deciding which cases are paradigmatic and important examples and which, thus, should be captured by a philosophical account, second, explicating assumptions that are only implicitly present in scientific practice, and third, constructing a coherent account, which involves assessing some claims of scientists as adequate, correct, or typical, and others as too vague or incorrect.

The more specific result of my meta-philosophical analysis was that my account of reduction must satisfy the following four criteria of adequacy: it should

(1*) (a) capture the *paradigmatic* and *important* cases of reduction that occur in current biological research practice,
(b) *explicate* the understanding of reduction that is (often only implicitly) present in biological practice,
(2) capture the *diversity* of the cases of reduction that are present in contemporary biology, but nevertheless be *coherent* and as *universal* as possible,
(3) clarify in which sense reductions contribute to or hinder the realization of relevant *epistemic norms*, for instance, explanatory success, and
(4) be *potentially useful* to contemporary biological practice (and to the broader society).[1]

In *Chapter II* I introduced the previous debate about reductionism in the philosophy of biology by pointing out four lessons that one should learn from it. These lessons are: first,

---

[1] In the following I will come back to these four criteria and show in which sense my account satisfies them.

before you discuss whether reductionism or antireductionism is true you should seek to understand what reduction is (Section 1). Second, if you seek to understand what reduction in actual biological practice is, you should be aware of the fact that epistemic issues rather than ontological issues are the ones that matter most (Section 2.3). But yet, you should pay attention to the respects in which epistemic and ontological issues are interrelated and also to the respects in which they are *not* interrelated (Section 2.2). Third, before you discuss epistemic reduction(ism) you need to specify which kind of epistemic reduction you are talking about, that is, whether the target of your analysis is theory reduction, methodological reduction, or explanatory reduction (Section 3). Fourth, it is time to move beyond the limitations of Nagel's classical model of reduction and, instead, to seek after a model of epistemic reduction, which is more adequate with respect to contemporary biological practice (Section 4).

Besides these four lessons I want to emphasize two further crucial results of *Chapter* II. The one concerns the notion of ontological reduction and the other the concept of a reductive method. First, in Section 2.1.1 I specified the notion of ontological reduction(ism) that is (explicitly or implicitly) taken for granted in philosophy of biology. My analysis disclosed that the notion of ontological reductionism is not – as it is in philosophy of mind – confined to type or property physicalism (which conceives reduction as a relation of identity between types, respectively between properties). Rather, what philosophers of biology commonly refer to as the kind of ontological reductionism that they presuppose in their discussions about epistemic reductionism is token physicalism. What is more, I argued that some philosophers of biology seem to spell out the relation of ontological reduction not as a relation of identity, but as a relation of constitution or of localization. However, it remains an open question what exactly the difference between constitution and identity is and whether the former really is a more adequate way to specify the notion of ontological reduction with respect to the biological sciences. These and other questions about ontological reduction are interesting questions which I could merely touch on in this book. They are worth being investigated in more detail in future philosophical work.

Second, in Section 3.2 I directed attention to the fact that in the current philosophical literature on methodological reductionism it is left unclear what exactly it means to practice biological research in a reductive fashion. I tried to fill this gap by specifying reductive methods as reductive strategies or heuristics (rather than as specific techniques or global stances) and by explicating paradigmatic and important reductive methods, such as the method of decomposition. This task was continued in Chapter V. There I identified two further reductive strategies that are implicitly discussed in biological research papers on reductionism, namely to ignore or to simplify the environmental factors of the system of interest (which implies conceptualizing contextual factors as background or as input conditions) and to study the parts of a system in isolation (which I specified as studying the parts of a system in other contexts than *in situ*).

In *Chapter III* I critically discussed the previous work that has been carried out on the topic of explanatory reduction in biology. More specifically, I examined two different perspectives on explanatory reduction, namely Rosenberg's view, according to which reduction is a relation between a higher-level and a lower-level explanation of the same phenomenon (Section 1), and Sarkar's, Hüttemann's, and Love's work on the reductive character of individual biological explanations (Section 2). The goal of this critical discussion was, on the one hand, to identify adequate and fruitful insights concerning explanatory reduction, which could then be utilized in developing my own analysis of explanatory reduction, and on the other hand, to sort out those ideas about explanatory reduction that have proven to be unconvincing. I concluded that Rosenberg's perspective encounters devastating criticism (Section 1.2 and 1.3). The most important objections are that his notion of explanatory reduction remains obscure (more precisely, what the result of the reduction process is), that his perspective on explanatory reduction is shaped too much by the philosophical requirements of his defense of explanatory reductionism, and that his account fails to capture actual cases of epistemic reduction from contemporary biological practice. These drawbacks lead me to dismiss Rosenberg's perspective and to choose the alternative path of focusing on individual reductive explanations and analyzing the constraints of their reductive character. I argued that explanatory reduction can be regarded as a relation between the two parts of an individual reductive explanation, namely between the description of the explanandum phenomenon and the description of the explanatory relevant factors (explanans). The central question that arose from this perspective was: under which conditions does this relation count as a reduction? In other words, which characteristics of biological explanations determine their reductive character? I pointed out that Sarkar, Hüttemann, and Love provide important insights into the reductivity of explanations, although they face criticism, too (Section 2.2.4 and 2.3.4). In short, they showed the close connection between reductive, part-whole, and mechanistic explanation. They argued in favor of a substantive, non-formal analysis of reductive explanation. And they proposed different criteria of reductivity or identified different aspects of reductive explanations, which constituted a fruitful point of reference in my own analysis of reductive explanation in the biological sciences.[2]

Chapter IV served to reveal the interrelations that exist between the issue of explanation and my account of explanatory reduction. Questions about explanation that are involved in disputes about explanatory reductionism were of interest in this chapter, too, but only insofar as they affect my analysis of reductive explanation. Among the several results of this chapter three findings are of particular importance and are worth being emphasized here. First, in Section 1.3 I argued that what makes an account of explanation ontic is *not* the thesis that explanations are objects or facts that exist in the world itself

---

[2] Since the insights that Sarkar, Hüttemann, and Love provide were picked up on and revised in my own account of explanatory reduction (Chapter V) I will not dwell on them here.

(rather than epistemic units like representations, models, or texts). Instead, I claimed that ontic accounts of explanation trace the difference between explanations and non-explanations, ultimately, back to differences in the world, and not to logical features of linguistic entities. That is, according to my view, the ontic character of a conception of explanation traces back to the thesis that what determines whether a description is explanatory is whether it truly represents certain causal facts in the world. In Chapter V this way to understand the term 'ontic' was applied in order to characterize my own account of explanatory reduction. Second, in Section 2.1 I pointed out that the question of what constitutes the reductive character of biological explanations (i.e. the question of explanatory reduction) does not boil down to the question of what characterizes an adequate explanation (i.e. the question of explanation). In other words, I endorsed the thesis that what makes an explanation reductive is different from what makes a description explanatory. In line with this, I argued that my account of explanatory reduction remains uncommitted with respect to whether CL or CM models of explanation adequately capture the nature of biological explanation and that this should be regarded as a virtue rather than as a shortcoming of my account. Third, in Section 2.2 I showed that the debate about explanatory *reductionism* (unlike the debate about explanatory *reduction*) boils down to answering specific questions about explanation, namely questions about the adequacy of higher- and lower-level explanations (i.e. whether certain levels of organization are explanatorily prior to others). I further revealed that how one answers these questions depends on one's stance on the pragmatics of explanation. I developed a revised version of van Fraassen's pragmatic account of explanation. According to my view, the research context determines which specific phenomenon *P* is to be explained and it affects according to which genuine relevance relation *R* certain information are included in the explanans and others are excluded. That is, in my view, explanation inevitably has pragmatic dimensions. However, I stressed that this does not imply that the adequacy of an explanation is "subjective" or exclusively determined by pragmatics. The reason that I adduced is that pragmatic factors neither influence which relevance relations count as genuine, nor do they influence how these relevance relations are to be specified. Taking this stance on the pragmatics of explanation is necessary since it ensures that all different kinds of adequate reductive explanations that can be found in actual biological practice are considered and are not dismissed in advance.

Let us now turn to the most significant findings of *Chapter V*, which constitutes the core of my book. This chapter contains my own analysis of the characteristics of reductive explanations in biology as well as of the merits and limitations of explaining biological phenomena in a reductive manner. Chapter V is divided into two main parts. In the *first part* (Section 1) I specified four concepts that are central to explicating the reductive character of biological explanations, namely the concept of a biological system and of its context or environment, the concept of decomposition and of a part, and the concept of levels of organization. These conceptual clarifications provided the basis on which, in the *second part* of this chapter (Sections 2 to 6), I could develop the core of my account of explanatory reduction by answering the central question of which features are typical for reductive explanations and allow distinguishing them from non-reductive explanations in biology. Before I summarize the answer that I gave to this question, let me recall how I specified the concept of a system, of parts, and of levels first.

In *Section 1.1* I argued that biological systems are composed of sets of interacting entities that form together a unified whole. This means that the parts of a biological system satisfy two conditions:

(1) they work together to bring about those behaviors that the system characteristically displays and
(2) they are located within a defined zone, which is surrounded by a spatial boundary.

I claimed that the spatial boundary that separates a system from its environment ideally is a continuous, constant surface, but that this is not always the case. The boundaries of some systems are subject to continuous change. Others are not continuous surfaces, but rather conglomerates of several disconnected boundaries. I revealed that it sometimes might be difficult to identify the boundary between a system and that biologists do not always agree on how to perform this task. However, I stressed that in most cases this does not imply that biologists *fail* to distinguish system and environment. Finally, I rejected the view that the separation of a system from its context is a highly flexible or even an arbitrary matter. I argued that, instead, the boundaries of many biological systems fall together with certain structures that exist in nature (e.g. with membranes), which is why I characterized them as bona fide boundaries, rather than as fiat boundaries. The question of how biological systems are individuated is closely related to the more specific question of how organisms are individuated or, to put it more generally, what an organism is and what distinguishes living things from non-living ones. These are major questions that lie at the heart of philosophy of biology, but that could only be touched on in this book. One prospect for future philosophical work in this context is to examine whether my account for individuating biological systems applies to organisms as well, and in which way it must be refined.[3]

In *Section 1.2* I analyzed what it means to decompose a system into its parts. I began with distinguishing three notions of parts:

(1) *mereological parts* (i.e. everything that counts as a part of a system according to CEM)
(2) *genuine parts* (i.e. any interacting entity that, first, contributes to bringing about at least one of the behaviors/functions a system typically exhibits and that, second, is located inside the spatial boundary of the system)

---

[3] One task will be to ascertain whether there can be identified just one behavior or one set of behaviors, according to which all organisms can be individuated (for instance, survival and reproduction).

(3) *components* (i.e. any interacting entity that is relevant to the behavior of a particular mechanism)

I concluded that with respect to biological systems the notion of a genuine part is the one that is most adequate. I argued that the notion of a mereological part is too unrestricted and that the notion of a component overlooks two important facts, namely that the parts of a system are individuated with respect to *all* behaviors that a system typically displays (and not only with respect to one) and that the parts of a system are located inside the spatial boundary of the system. Moreover, I rejected the pluralistic view that different theoretical perspectives or partitioning frames result in different, equally adequate, non-coincident, and (partly) inconsistent decompositions of the same system. I adopted a more optimistic view and stated that the different decompositions of a system frequently are consistent and can be integrated via other compositional relations. Questions about how to individuate the parts of a system (as well as the systems themselves) and in which way one's theoretical perspective affects this individuation give rise to more general questions about the relation between scientific concepts and the natural world. For instance, do scientific concepts "carve nature at its joints", do they refer to so called natural kinds, or do they "merely" reflect our (or the biologist's) way of classifying the world? Although it might be that some of my theses include some commitments concerning how to answer these questions, it lies outside of the scope of this book to argue in favor of a certain position in this debate. To make explicit these commitments and to underpin my account with a clear stance on natural kinds and on related issues constitutes an important task for future philosophical work.

Section 1.3 served to clarify the notion of levels of organization that my analysis of reductive explanation presupposes. The result of my critical examination of different accounts of levels is that, first, levels of science (i.e. of theories) are a deceptive indicator for levels of nature since different theories from different disciplines refer to entities at the same level and since often one theory appeals to entities from different levels. Second, I argued that size can be no more than a rough indicator of levels, too, because size relations are only by-products of compositional relations, which are the proper relations that determine levels. Third, I emphasized, however, that it must be specified *which* part-whole relations give rise to different levels of organization. I showed that mereological-part-system relations are inadequate, just as Craver's component-mechanism relations. I claimed that, instead, genuine-part-system relations seem to be the ones that give rise to a convincing notion of levels of nature. Finally, I pointed out that the notion of level that underlies discussions about reductionism and reductive explanation in biology cannot be captured by a concept of locally defined levels. Rather, I argued that my analysis needs an account of global levels. I assumed that such a global notion of level may be obtained, for example, by bringing together different genuine-part-system distinctions and establishing a hierarchy of different genuine-part-system relations. However, it remains a significant task for future philosophical work to fill in the details of such a global notion of levels and to reveal its adequacy not only for the context of reductive explanation, but also for the biological sciences in general.

I concluded that the investigation of biological systems and the development of explanations for the behaviors of these systems involve three major tasks: first, the individuation of the system of concern (which includes the separation of the system from its environment), second, the decomposition of that system into its genuine parts, and third, the identification of different levels of organization and the association of certain interacting entities to these levels. The following figure illustrates this thesis:

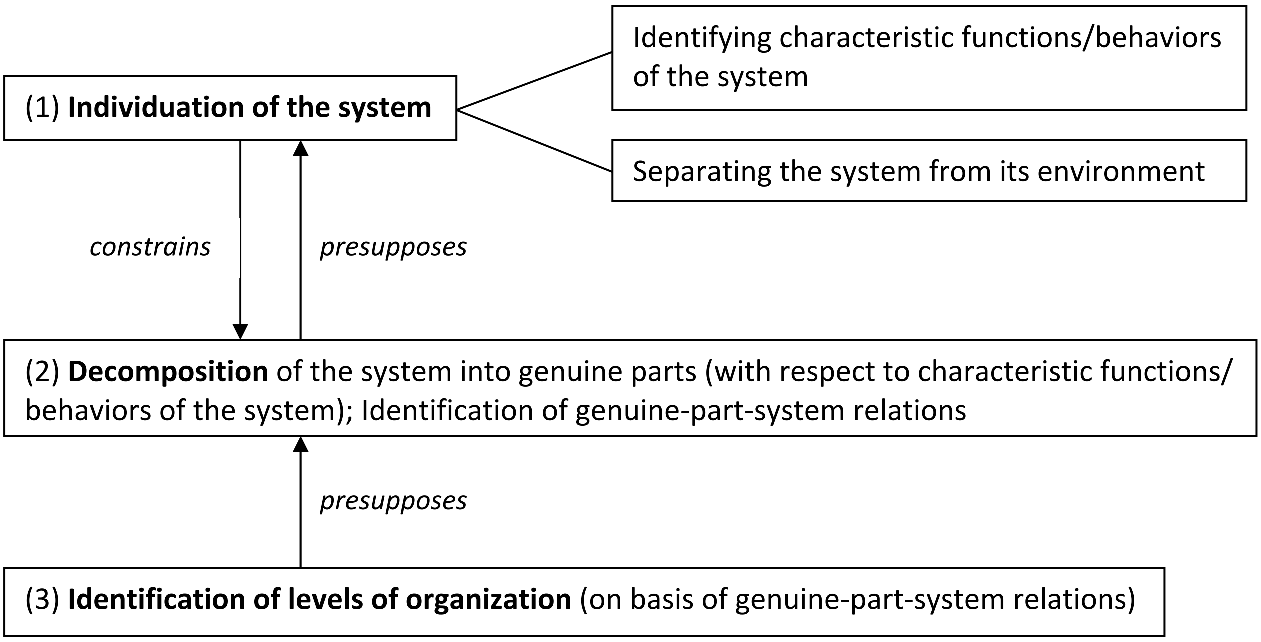

Figure V.11: The Three Tasks of Investigating Biological Systems

Based on these conceptual clarifications I could then in the *second part* of Chapter V (Section 2 to 6) address the central question of my book, namely which characteristics are common to reductive explanations in the biological science, and distinguish them from non-reductive explanations. I argued that reductive explanations exhibit two necessary features. They explain the behavior of a biological system by referring only to

(1) factors that are located on a lower level of organization than the system, and to
(2) parts in isolation (i.e. to those relational properties of and interactions between parts that can be studied in other contexts than *in situ*).

I identified a third characteristic that many (but not all) reductive explanations exhibit. They explain the behavior of a biological system by focusing on

(3) factors that are internal to the system (i.e. that are genuine parts of the system).

I claimed that there are many reasons in favor of the view that, taken together, these three conditions are sufficient for a biological explanation to be reductive. Let us recall how I specified these three characteristics of reductive explanations in more detail.

In *Section 2* I specified the lower-level character of reductive explanations in two ways: on the one hand, I pointed out that reductive explanations are characterized by a unidirectional flow of explanation from the lower to the higher level, which is why the appeal to downward causation renders an explanation non-reductive. On the other hand, I clarified what exactly it means for an explanation to refer solely to lower-level factors. My analysis yielded three main results: first, I argued that the set of lower-level factors is not confined to the set of genuine parts of the particular system of interest. Rather, it comprises also factors that are external to the system, but that are located on the same level(s) as the genuine parts of the system. Second, I claimed that the category of lower-level factors encompasses functional properties of genuine parts, too, and that reference to functional properties of parts does not violate the reductive character of an explanation as long as the context that needs to be included is described in lower-level terms, too. Third, I stated that the organization of the system's parts counts as a lower-level factor if it can be characterized as a lower-level relational property (i.e. as a property of lower-level objects) and not as a systemic property (i.e. as a property of the system).

A second major result of Section 2 was the identification of two subtypes of lower-level explanation, namely fundamental-level explanation and single-factor explanation. I characterized fundamental-level explanations as reductive explanations that refer only to factors that are located on the lowest, the fundamental biological level of molecules (i.e. they are molecular explanations). My main thesis was that fundamental-level explanation constitutes an important subtype of reductive explanations, but that philosophers should not be fooled into identifying reductive explanation with fundamental-level explanation. I argued that doing so obscures the diversity of reductive explanation in biological science and converts explanatory reductionism into a straw man. I characterized single-factor explanations as such explanations of the behavior of a system that refer only to a single causal factor. I argued that reductive explanations of this kind are inadequate with respect to most biological phenomena, which is why most biologists refer to them in a negative way (e.g. when they reject reductionism).

Section 3 introduced a second feature that I regard as being typical (but not necessary) for the reductive character of biological explanations. In my view, many reductive explanations explain the behavior of a system by focusing on those factors that are internal to this system. I argued that any factor that is internal to a particular system is a genuine part of this system, which is why I refer to explanations with an internal character as part-whole explanations. Because the internal character of reductive explanations has been overlooked or intermingled with their lower-level character for a long time I claimed that it is important to be aware of the differences between these two features of reductive explanations. The result of my analysis was that the category of lower-level factors is broader than the category of internal factors. In other words, I argued that all explanations with an internal character are lower-level explanations, but that the set of lower-level explanations also include reductive explanations that fail to possess an internal character. Finally, I elaborated upon the thesis that reductive explanation focus on internal factors. My main thesis was that this implies that either environmental factors are ignored altogether (i.e. not mentioned in the explanation at all) or that they are simplified to a great extent. I argued that the context of a system can be simplified in two major ways: on the one hand, environmental factors can be simplified as background conditions, which are treated as being homogeneous or remaining constant over time. On the other hand, they can be simplified as mere input conditions, which need to be satisfied only at the beginning of the process that brings about the explanandum phenomenon "automatically" and whose specific character and origin is regarded as irrelevant. All in all, my thesis that the internal character is a typical but not a necessary condition for the reductivity of an explanation leaves room for the assertion that evolutionary or ecological explanations, which often include the description of environmental factors as an essential part of the explanantia, can be reductive, too. For instance, evolutionary explanations may be reductive because they appeal to selection processes that involve only organisms (rather than groups) or only genes (rather than organisms), or because they are mono-causal explanations of evolution, that is, adaptationist explanations that ignore other factors for generating evolutionary change aside from natural selection (e.g. genetic drift, developmental constraints, etc.). However, future philosophical work must show whether my account of explanatory reduction succeeds in capturing *all* kinds of reductive explanations that are present in different biological fields – for example, whether it sufficiently accounts for what reductive explanations in evolutionary biology are.[4] A related prospect for future philosophical work will be to spell out the exact relation between reductive and functional explanation. My assumption is that, even though functional explanations of the behavior of a biological system inevitably appeal to contextual factors, they can still be reductive if these contextual factors are described in lower-level terms.

In *Section 4* I presented a third characteristic of reductive explanations in biology, namely that they explain the behavior of a system by appealing exclusively to the system's parts in isolation. I argued that philosophers have not paid much attention to this necessary feature of reductive explanations yet, but that the current vivid discussions in the biological literature about the limits of reductionism demonstrate that they should. I argued that the phrase 'to study parts in isolation' does not mean that the genuine parts of a system are investigated completely on their own, but rather that the parts are studied in isolation from their original system, that is, in other contexts than *in situ*. I further elaborated this thesis by pointing out that studying parts in isolation does not result in reductive explanations, which refer only to intrinsic properties of the system's parts. Rather, I claimed that it results in reductive explanations that appeal only to those relational properties of parts and to those interactions among them that can be discovered by investigating the parts in contexts other

---

[4] For more details see for instance Weber 1998.

than *in situ*. Finally, I argued that this third feature of reductive explanations can be rephrased as describing the genuine parts of a system as being organized in a simple way, as interacting with each other in a sequential and linear fashion, and as exhibiting properties that are largely unaffected by their (intrasystemic) context.

In *Section 5* I clarified the similarities and differences that exist among three kinds of explanation, namely between reductive, part-whole, and mechanistic explanation. The main results are summarized in the following table:

	reductive explanation	**part-whole explanation**	**mechanistic explanation**
lower-level character	necessary	not necessary	necessary
internal character	not necessary	necessary	necessary
parts-in-isolation character	necessary	not necessary	not necessary

Table V.25: Comparing Reductive, Part-whole, and Mechanistic Explanations

At the end of this chapter, in *Section 6*, I explained why I call my account of explanatory reduction “ontic”. I pointed out that this denotes that I trace back the difference between reductive explanations and non-reductive explanations to particular relations that exist in the world, rather than to certain logical relations between explanandum-sentence and explanans-sentence. In other words, according to my account, what determines whether an explanation is reductive or not is whether or not it truly represents specific relations or facts that exist in the world, namely whether the explanatorily relevant factors are described as being located on a lower level than the system whose behavior is to be explained, whether they are represented as being internal to the system, and whether they are described as being parts in isolation.

Finally, let me come back to the criteria of adequacy that I identified in Chapter I and show in which sense my account of explanatory reduction in biology satisfies them. The three characteristics of reductive explanation that I identify are the result of an extensive analysis, on the one hand, of paradigmatic and important examples of reductive explanations from contemporary biological practice, such as the molecular explanation of photosynthesis, the genetic explanation of carcinogenesis, or the explanation of food web structure by IBMs (criterion 1*a), and on the other hand, of discussions about reductionism that can be found in current biological research papers. In particular, the analysis of the biological literature frequently involved making explicit assumptions about reductive explanations that are only implicitly present in the literature (criterion 1*b). This is due to the fact that biologists primarily are interested in the appropriateness of applying certain reductive methods or in the adequacy of certain explanations. By contrast, my central question of what it is that makes biological explanations reductive and distinguishes them from non-reductive ones is not important to them by itself. The empirical data that I took into account came from various different biological fields (ranging from molecular biology and developmental biology over ecology to the biomedical science). Accordingly, in developing my account of explanatory reduction I tried to capture much of the diversity of reductive reasoning that is present in different biological disciplines. However, my aim was also to reconstruct a coherent philosophical account which is as universal as possible (criterion 2). Establishing coherency involved assessing certain assumptions of biologists as incorrect or as inadequate, for instance, the view that to explain a phenomenon reductively amounts to treating the system whose behavior is to be explained as an aggregative system or the thesis that reductive explanations disregard all relational properties of and interactions between a system’s parts. The central goal of developing my account of explanatory reduction was to find an answer to the question of what determines the reductivity of a biological explanation and distinguishes it from a non-reductive explanation. This aim can be reformulated as seeking the norms of reductive explanation, that is, the standards according to which an explanation is judged as being reductive or non-reductive. However, the normative character of my account results from a different feature, namely from the fact that it clarifies in which sense reductions contribute to or hinder the realization of a relevant epistemic norm, that is, explanatory success (criterion 3). More specifically, by identifying the three characteristics of reductive explanations I also reveal what the merits and limitations of reductive explanations are, that is, under which conditions they are adequate and under which conditions they fail to be adequate. For instance, reductive explanations are inadequate if reference to higher-level factors is necessary for explaining the behavior of a system or if the system whose behavior is to be explained is only minimally decomposable or not decomposable at all, rather than being nearly decomposable. Since my account is sensitive to contemporary biological practice it also has the potential to be useful to biologists (criterion 4). Most notably, my analysis clarifies the concept of a reductive explanation (and of a reductive method) and it specifies what the strengths and limits of reductive explanations are. That way, my account has the potential to enhance debates about explanatory reductionism – not only in philosophy, but also in biological science itself.

References

Achinstein, Peter (1983): *The Nature of Explanation*, New York: Oxford University Press.

Ahn, Andrew C./ Tewari, Muneesh/ Poon, Chi-Sang/ Phillips, Russell S. (2006a): "The Limits of Reductionism in Medicine: Could Systems Biology Offer an Alternative?", in: *PLoS Medicine* 3 (6): 709-713.

Ahn, Andrew C./ Tewari, Muneesh/ Poon, Chi-Sang/ Phillips, Russell S. (2006b): "The Clinical Applications of a Systems Approach", in: PLoS *Medicine* 3 (7): 1-5.

Alberts, Bruce/ Johnson, Alexander/ Lewis, Julian/ Raff, Martin/ Roberts, Keith/ Walter, Peter (2008): *Molecular Biology of the Cell*. 5th Edition, New York: Garland Science.

Ayala, Francisco J. (1974): "Introduction", in: Ayala, Francisco J./ Dobzhansky, Theodosius (eds.): *Studies in the Philosophy of Biology,* Berkeley: University of California Press, vii-xvi.

Baker, Lynne Rudder (1997): "Why Constitution in not Identity", in: *Journal of Philosophy* 94 (12): 599-621.

Baker, Lynne Rudder (2000): *Persons and Bodies: A Constitution View*, New York: Cambridge University Press.

Baker, Lynne Rudder (2007): *The Metaphysics of Everyday Life*. An Essay in Practical Realism, Cambridge: Cambridge University Press.

Baker, Lynne Rudder (2009): "Non-reductive Materialism", in: McLaughlin, Brian P. (ed.): *The Oxford Handbook of Philosophy of Mind*, Oxford: Clarendon Press, 109-127.

Bateson, Patrick (2005): "The Return of the Whole Organism", in: *Journal of Bioscience* 30 (1): 31-39.

Beatty, John (1981): "What's Wrong With the Received View of Evolutionary Theory?", in: *PSA* 1980 (2): 397-426.

Beatty, John (1990): "Evolutionary Anti-Reductionism: Historical Considerations", in: *Biology and Philosophy* 5: 199-210.

Beatty, John (1995): "The Evolutionary Contingency Thesis", in: Wolters, Gereon/ Lennox, James (eds.): *Concepts, Theories, and Rationality in the Biological Sciences*. Pittsburgh: University of Pittsburgh Press, 45-81.

Bechtel, William (2001): "The Compatibility of Complex Systems and Reduction: A Case Analysis of Memory Research", in: *Minds and Machines* 11: 483-502.

Bechtel, William (2006): *Discovering Cell Mechanisms.* The Creation of Modern Cell Biology, Cambridge: Cambridge University Press.

Bechtel, William (2008): "Mechanisms in Cognitive Psychology: What Are the Operations?", in: *Philosophy of Science* 75: 983-994.

Bechtel, William (2008): *Mental Mechanisms*. Philosophical Perspectives on Cognitive Neuroscience, New York/ London: Taylor and Francis Group.

Bechtel, William (2009): “Looking down, around, and up: Mechanistic explanation in psychology”, in: *Philosophical Psychology* 22 (5): 543-564.

Bechtel, William/ Abrahamsen, Adele (2005): “Explanation: A Mechanist Alternative”, in: *Studies in History and Philosophy of Biological and Biomedical Sciences* 36: 421-441.

Bechtel, William/ Abrahamsen, Adele (2011): “Complex Biological Mechanisms: Cyclic, Oscillatory, and Autonomous”, in: Hooker, Cliff A. (ed.): *Philosophy of Complex Systems. Handbook of the Philosophy of Science*, New York: Elsevier, 257-285.

Bechtel, William/ Hamilton, Andrew (2007): “Reduction, Integration, and the Unity of Science: Natural, Behavioral, and Social Sciences and the Humanities”, in: Kuipers, Theo A. F. (ed.): *General Philosophy of Science: Focal Issues.* Amsterdam: Elsevier, 377-430.

Bechtel, William/ Richardson, Robert C. (2010): *Discovering Complexity*. Decomposition and Localization as Strategies in Scientific Research, Cambridge: MIT Press.

Beckermann, Ansgar (1992): “Introduction - Reductive and Nonreductive Physicalism”, in: Beckermann, Ansgar/ Flohr, Hans/ Kim, Jaegwon (eds.): *Emergence or Reduction?* Essays on the Prospects of Nonreductive Physicalism, Berlin: de Gruyter, 1-21.

Bedau, Mark A./ Humphreys, Paul (2008): *Emergence.* Contemporary Readings in Philosophy and Science, Cambridge: MIT Press.

Beresford, Mark J. (2010): “Medical Reductionism: Lessons from the Great Philosophers”, in: *Q J Med* 103 (9): 721-724.

Bickle, John (1998): *Psychoneural Reduction.* The New Wave, Cambridge: MIT Press.

Bickle, John (2003): *Philosophy and Neuroscience.* A Ruthlessly Reductive Account, Dordrecht: Kluwer Academic Publishers.

Bickle, John (2006): “Reducing Mind to Molecular Pathways: Explicating the Reductionism Implicit in Current Cellular and Molecular Neuroscience”, in: *Synthese* 151: 411-434.

Bizzarri, Mariano/ Cucina, Alessandra/ Conti, Filippo/ D’Anselmi, Fabrizio (2008): “Beyond the Oncogene Paradigm: Understanding Complexity in Carcinogenesis”, in: *Acta Biotheoretica* 56: 173-196.

Bock, Gregory R./ Goode, Jamie A. (1998) *The Limits of Reductionism in Biology.* Novartis Foundation Symposium no. 213, Chichester: Wiley.

Boyd, Richard (1980): “Materialism without Reductionism: What Physicalism Does Not Entail”, in: Block, Ned J. (ed.): *Readings in Philosophy of Psychology*, vol. 1, Cambridge: Harvard University Press, 67-106.

Boyle, Robert (1966): “The Excellency of Theology (Works IV)”, in: *The Works of the Honourable Robert Boyle*. In Six Volumes. To Which Is Prefixed the Life of the Author. A New Edition, London: Printed for J. and F. Rivington et al.

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

Brandon, Robert N. (1996): *Concepts and Methods in Evolutionary Biology.* Cambridge: Cambridge University Press.

Brigandt, Ingo (forthcoming): “Explanation in Biology: Reduction, Pluralism, and Explanatory Aims”, in: *Science and Education*. Online first, March 2011.

Brigandt, Ingo/ Love, Alan (2008): “Reductionism in Biology”, in: Zalta, Edward N. (ed.): *The Stanford Encyclopedia of Philosophy.* (Fall 2008 Edition), URL = <http://plato.stanford.edu/ archives/fall2008/entries/reduction-biology/>.

Broad, Charlie D. (1925): *The Mind and its Place in Nature*, London: Paul/ Routledge.

Buchanan, Anne V./ Weiss, Kenneth M./ Fullerton, Stephanie M. (2006): “Dissecting Complex Disease: the Quest for the Philosopher’s Stone?”, in: *International Journal of Epidemiology* 35: 562-571.

Byerly, Henry (2003): “Reductionism: Analysis and Synthesis in Biological Explanations”, in: *The Quarterly Review of Biology* 78 (3):336-342.

Callebaut, Werner (1993): *Taking the Naturalism, or, How Real Philosophy of Science Is Done,* Chicago/ London: University of Chicago Press.

Campbell, Neil A./ Reece, Jane B./ Urry, Lisa A./ Cain, Michael L./ Wasserman, Steven A./ Minorsky, Peter V./ Jackson, Robert B. (2005): *Biology.* 7th ed., San Francisco: Pearson/ Cummings.

Carnap, Rudolf (1950): *Logical Foundations of Probability*, Chicago: University of Chicago Press.

Carrier, Martin (2007): “Wege der Wissenschaftsphilosophie im 20. Jahrhundert”, in: Bartels, Andreas/ Stöckler, Manfred (eds.): *Wissenschaftstheorie*. Ein Studienbuch, Paderborn: mentis, 15-44.

Carruthers, Peter (2004): “Reductive Explanation and the ‘Explanatory Gap’”, in: *Canadian Journal of Philosophy* 34: 153-174.

Cartwright, Nancy (1999): *The Dappled World: A Study of the Boundaries of Science*, Cambridge: Cambridge University Press.

Casadevall, Arturo/ Fang, Ferric C./ Pirofski, Liise-Anne (2011): “Microbial Virulence as an Emergent Property: Consequences and Opportunities”, in: *PLoS Pathogens* 7: 1-3.

Chalmers, David J. (1996): *The Conscious Mind*. In Search of a Fundamental Theory, New York/ Oxford: Oxford University Press.

Chalmers, David J./ Jackson, Frank (2001): “Conceptual Analysis and Reductive Explanation”, in: *Philosophical Review* 110 (3): 315-360.

Chong, Lisa/ Ray, L. Bryan (2002): “Whole-istic Biology”, in: *Science* 295: 1661.

Coffa, J. Alberto (1974): “Hempel's Ambiguity”, in: *Synthese* 28 (2): 141 – 163.

Craver, Carl F. (2002a): “Structures of Scientific Theories”, in: Machamer, Peter/ Silberstein, Michael (eds.): *The Blackwell Guide to the Philosophy of Science*. Malden/ Oxford: Blackwell Publishers, 55-79.

Craver, Carl F. (2002b): “Interlevel Experiments and Multilevel Mechanisms in the Neuroscience of Memory”, in: *Philosophy of Science* 69: 83-97.

Craver, Carl F. (2005): “Beyond Reduction: Mechanisms, Multifield Integration, and the Unity of Neuroscience”, in: *Studies in the History and Philosophy of Biological and Biomedical Sciences* 36: 373-395.

Craver, Carl F. (2006): “When Mechanistic Models Explain”, in: *Synthese* 153: 355-376.

Craver, Carl F. (2007a): *Explaining the Brain.* Mechanisms and the Mosaic Unity of Neuroscience, Oxford: Oxford University Press.

Craver, Carl F. (2007b): “Constitutive Explanatory Relevance”, in: *Journal of Philosophical Research* 32: 3-20.

Craver, Carl F. (2008): “Physical Law and Mechanistic Explanation in the Hodgkin and Huxley Model of the Action Potential”, in: *Philosophy of Science* 75: 1022-1033.

Craver, Carl F. (forthcoming a): “The Ontic Conception of Scientific Explanation”, in: Hüttemann, Andreas/ Kaiser, Marie I./ Scholz, Oliver (eds.): *Explanation in the Special Sciences – The Case of Biology and History* (Synthese Library), Dordrecht: Springer.

Craver, Carl F. (forthcoming b): “Functions and Mechanisms: A Perspectivalist Account”, in: Huneman, Philippe (ed.): *Functions: Selection and Mechanisms* (Synthese Library), Dordrecht: Springer.

Craver, Carl F./ Bechtel, William (2007): “Top-down Causation without Top-down Causes”, in: *Biology and Philosophy* 22: 547-563.

Craver, Carl F./ Darden, Lindley (2005): “Introduction”, in: *Studies in History and Philosophy of Biological and Biomedical Sciences* 36: 233-244.

Craver, Carl F./ Darden, Lindley. (2001): “Discovering Mechanisms in Neurobiology: The Case of Spatial Memory”, in: Machamer, Peter/ Grush, Rick/ McLaughlin, Peter (eds.): *Theory and Method in Neuroscience.* Pittsburgh: University of Pittsburgh Press, 112-137.

Crick, Francis H. C. (1966): *Of Molecules and Men*. Seattle: University of Washington Press.

Crick, Francis H. C. (1988): *What Mad Pursuit.* A Personal View of Scientific Discovery, New York: Francis Books.

Cummins, Robert (1975): “Functional Analysis”, in: *The Journal of Philosophy* 72 (20): 741-765.

Darden, Lindley (1991): *Theory Change in Science.* Strategies from Mendelian Genetics. New York: Oxford University Press.

Darden, Lindley (2005): “Relations among Fields: Mendelian, Cytological and Molecular Mechanisms”, in: *Studies in History and Philosophy of Biological and Biomedical Sciences* 36: 357-371.

Darden, Lindley (2008): “Thinking Again about Biological Mechanisms”, in: *Philosophy of Science* 75: 958-969.

Darden, Lindley/ Maull, Nancy (1977): “Interfield Theories”, in: *Philosophy of Science* 44, 43-64.

Darden, Lindley/ Tabery, James (2010): “Molecular Biology”, in: Zalta, Edward N. (ed.): *The Stanford Encyclopedia of Philosophy* (Fall 2010 Edition), URL = <http://plato.stanford.edu/archives/fall2010/entries/molecular-biology/>.

Davidson, Donald (1970): “Mental Events”, in: Foster, Lawrence/ Swanson, J. W. (eds.): *Experience and Theory*, London: Duckworth, 79-101.

Dawkins, Richard (1976): *The Selfish Gene*. New York: Oxford University Press.

Dawkins, Richard (1982): *The Extended Phenotype: the Gene as the Unit of Selection*, Oxford: Freeman.

De Vreese, Leen/ Weber, Erik/ Van Bouwel, Jeroen (2010): “Explanatory Pluralism in the Medical Science: Theory and Practice”, in: *Theoretical Medicine and Bioethics* 31: 371-390.

Delehanty, Megan (2005): “Emergent Properties and the Context Objection to Reduction”, in: *Biology and Philosophy* 20: 715-734.

Dennett, Daniel C. (2005): *Darwin’s Dangerous Idea*. New York: Simon and Schuster.

Descartes, René (1991 [1644]): *Principles of Philosophy*. Translated, with Explanatory Notes, by V. R. Miller and R. P. Miller, Dordrecht: Kluwer.

Dizadji-Bahmani, Foad/ Frigg, Roman/ Hartmann, Stephan (2010): “Who Is Afraid of Nagelian Reduction?”, in: *Erkenntnis* 73: 393-412.

Dobson, Christopher M. (2003): “Protein Folding and Misfolding”, in: *Nature* 426: 884-890.

Dowe, Phil (1992): “Process Causality and Asymmetry”, in: *Erkenntnis* 37 (2): 179-196.

Dowe, Phil (2000): *Physical Causation*, Cambridge: Cambridge University Press.

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

Dupré, John (2009): “It Is Not Possible to Reduce Biological Explanations to Explanations in Chemistry and/ or Physics”, in: Ayala, Francisco J./ Arp, Robert (eds.): *Contemporary Debates in Philosophy of Biology,* Chichester: Blackwell, 32-47.

Dupré, John (2010). “The Polygenomic Organism”, in: *Sociological Review* 58 (1): 19-31.

Einstein, Albert (1933): On the Method of Theoretical Physics. The Herbert Spencer Lecture, delivered at Oxford, 10 June 1933, Oxford: Clarendon Press.

Evnine, Simon J. (2011): “Constitution and Composition: Three Approaches to their Relation”, in: *ProtoSociology* 27: 212-235.

Falkenburg, Brigitte (2005): “Der Wert wertfreier Wissenschaft“, in: Gesang, Bernward (ed.): *Deskriptive oder normative Wissenschaftstheorie?* Frankfurt: ontos, 91-122.

Fang, Ferric C./ Casadevall, Arturo (2011): “Reductionistic and Holistic Science”, in: *Infection and Immunity* 79: 1401-1404.

Fazekas, Peter /Kertész, Gergely (2011): “Causation at Different Levels: Tracking the Commitments of Mechanistic Explanations”, in: *Biology and Philosophy* 26: 365-383.

Feigl, Herbert, (1958): “The ‘Mental’ and the ‘Physical’”, in: Feigl, Herbert/ Scriven, Michael/ Maxwell, Grover (eds.): *Concepts, Theories and the Mind-Body Problem*. Minnesota Studies in the Philosophy of Science, vol. 2, Minneapolis: University of Minnesota Press.

Feyerabend, Paul (1962): “Explanation, Reduction and Empiricism”, in: Feigl, Herbert / Maxwell, Grover (eds.): *Scientific Explanation, Space, and Time*. Minneapolis: University of Minnesota Press, 28-97.

Fincham, John Robert Stanley (2000): “Reductionism should be clarified, not dismissed”, in: *Nature* 406: 343.

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

Folse 3rd, Henri. J./ Roughgarden, J. (2010): “What is an Individual Organism? A Multilevel Selection Perspective”, in: *The Quarterly Review of Biology* 85 (4): 447-472.

Forge, John (1998): “Explanation and Mechanism: Reflections on the Ontic Conception of Explanation”, in: Anapolitanos, Dionysios/ Baltas, Aristides/ Tsinorema, Stavroula (eds.): *Philosophy and the Many Faces of Science.* Lanham: Rowman & Littelfield Publishers, 76-92.

Friedman, Michael (1974): “Explanation and Scientific Understanding”, in: *Journal of Philosophy* 71 (1): 5-19.

Frost-Arnold, Greg (2004): “How to Be an Anti-reductionist about Developmental Biology: Response to Laubichler and Wagner”, in: *Biology and Philosophy* 19: 75-91.

Frydman, Judith (2001): “Folding of Newly Translated Proteins *in vitro*: The Role of Molecular Chaperones”, in: *Annual Review of Biochemistry* 70: 603-647.

Gallagher, Richard/ Appenzeller, Tim (1999): “Beyond Reductionism”, in: *Science* 284: 79.

Gatherer, Derek (2010): “So what do we really mean when we say that systems biology is holistic?”, in: *BMC Systems Biology* 4: 1-12.

Giere, Ronald N. (1988): *Explaining Science.* A Cognitive Approach, Chicago: University of Chicago Press.

Giere, Ronald N. (1999): *Science without Laws*, Chicago/ London: University of Chicago Press.

Gilbert, Scott F. (2006): *Developmental Biology*. 8th Edition, Sunderland: Sinauer Associates.

Gillett, Carl (2007): “Understanding the Reductionism: The Metaphysics of Science and Compositional Reduction”, in: *Journal of Philosophy* 104 (4): 193-216.

Glennan, Stuart S. (2008): “Mechanism”, in: Psillos, Stathis/ Curd, Martin (eds.): *The Routledge Companion to Philosophy of Science.* London/ New York: Routledge, 376-384.

Glennan, Stuart S. (2010): “Ephemeral Mechanisms and Historical Explanation”, in: *Erkenntnis* 72 (2): 251-266.

Glennan, Stuart S. (2002): “Rethinking Mechanistic Explanation”, in: *Philosophy of Science* 69: 342-353.

Godfrey-Smith, Peter (2008): “Reduction in Real Life”, in: Hohwy, Jakob/ Kallestrup, Jesper (eds.): *Being Reduced.* New Essays on Reduction, Explanation and Causation, Oxford: Oxford University Press, 52-74.

Greenspan, Ralph J. (2001): “The Flexible Genome”, in: *Nature Reviews* 2: 383- 387.

Grene, Marjorie (1987): “Hierarchies in Biology”, in: *American Scientist* 75: 504-510.

Griffiths, Paul E. (2007): “Philosophy of Biology”, in: Sarkar, Sahotra/ Pfeifer, Jessica (eds.): *The Philosophy of Science*. An Encyclopedia, New York: Routledge, 68-75.

Grizzi, Fabio/ Chiriva-Internati, Maurizio (2005): “The Complexity of Anatomical Systems”, in: *Theoretical Biology and Medical Modelling* 2: 26-34.

Grizzi, Fabio/ Chiriva-Internati, Maurizio (2006): “Cancer: Looking for Simplicity and Finding Complexity”, in: *Cancer Cell International* 6: 4-10.

Grizzi, Fabio/ Di Ieva, Antonio/ Russo, Carlo /Frezza, Eldo E./ Cobos, Everardo/ Muzzio, Pier Carlo/ Chiriva-Internati, Maurizio (2006): “Cancer initiation and progression: an unsimplifiable complexity”, in: *Theoretical Biology and Medical Modelling* 3: 37-42.

Hahn, William C./ Weinberg, Robert A. (2002): “Modelling the Molecular Circuitry of Cancer” in: *Nature Reviews* 2: 331-341.

Hartl, F. Ulrich/ Hayer-Hartl, Manajit (2002): “Molecular Chaperones in the Cytosol: from Nascent Chain to Folded Protein”, in: *Science* 295: 1852-1858.

Hempel, Carl G. (1965): *Aspects of Scientific Explanation*. And Other Essays in the Philosophy of Science, New York: Free.

Hempel, Carl G./ Oppenheim, Paul (1848): “Studies in the Logic of Explanation”, in: *Philosophy of Science* 15 (2): 135-175.

Hitchcock, Christopher (1995): “Discussion: Salmon on Explanatory Relevance”, in: *Philosophy of Science* 62: 304-20.

Hitchcock, Christopher/ Woodward, James (2003): “Explanatory Generalizations, Part II: Plumbing Explanatory Depth”, in: *Noûs* 37 (2): 181-199.

Hooker, Clifford (1981): “Towards a General Theory of Reduction. Part I: Historical and Scientific Setting. Part II: Identity in Reduction. Part III: Cross-Categorical Reduction”, in: *Dialogue* 20: 38-59, 201-236, 496-529.

Hove, J. R./ Köste, R. W./ Forouhar, A. S./ Acevedo-Bolton, G./ Fraser, S. E./ Gharib, M. (2003): “Intracardiac Fluid Forces are an Essential Epigenetic Factor for Embryonic Cardiogenesis”, in: *Nature* 421: 172-177.

Hoyningen-Huene, Paul (2007): “Reduktion und Emergenz“, in: Bartels, A./ Stöckler, M. (eds.): Wissenschaftstheorie. Ein Studienbuch, Paderborn: mentis, 177-197.

Hull, David (1972): “Reductionism in Genetics – Biology or Philosophy?”, in: *Philosophy of Science* 39: 491-499.

Hull, David (1974): *The Philosophy of Biological Science*, New Jersey: Prentice-Hall Inc.

Hull, David (1976): “Informal Aspects of Theory Reduction”, in: *PSA* 1974, 653-670.

Hull, David L./ van Regenmortel, Marc H. V. (2002): “Introduction”, in: van Regenmortel, Marc H. V./ Hull, David L. (eds.): *Reductionism in the Biomedical Sciences*, London: John Wiley & Sons, 1-13.

Huneman, Philippe (2010): “Topological explanations and robustness in biological sciences“, in: *Synthese* 177: 213-245.

Hunter, Philip (2003): “Putting Humpty Dumpty Back Together Again”, in: *Scientist* 17 (4): 20-22.

Hüttemann, Andreas (2004): *What’s Wrong With Microphysicalism?* New York: Routledge.

Hüttemann, Andreas/ Love, Alan C. (2011): “Aspects of Reductive Explanation in Biological Science: Intrinsicality, Fundamentality, and Temporality”, in: *British Journal for Philosophy of Science* 62 (3): 519-549.

Huxley, A. F./ Niedergerke R. (1954): “Structural Changes in Muscle During Contraction; Interference Microscopy of Living Muscle Fibres”, in: *Nature* 1973: 971-973.

Huxley, H. E./ Hanson J.(1954): “Changes in the Cross-Striations of Muscle During Contraction and Stretch and Their Structural Interpretation”, in: *Nature* 1973: 973-976.

Jackson, Frank (1998): *From Metaphysics to Ethics.* A Defense of Conceptual Analysis, Oxford: Oxford University Press.

Janich, Peter (2005): “Wissenschaftsphilosophie als kritische Reflexion auf eine historische Praxis“, in: Gesang, Bernward (ed.): *Deskriptive oder normative Wissenschaftstheorie?* Frankfurt: ontos, 145-166.

Johnston, Mark (1992): “Constitution Is Not Identity”, in: *Mind* 101 (401): 89-106.

Johnston, Mark (2006): “Hylomorphism”, in: *Journal of Philosophy* 103 (12):652-698.

Joyner, Michael J./ Pedersen, Bente K. (2011): “Ten questions about systems biology“, in: *The Journal of Physiology* 589 (5): 1017-1030.

Kauffmann, Stuart A. (1970): “Articulation of Parts Explanation in Biology and the Rational Search for Them”, in: *PSA* 1970: 257-272.

Keil, Geert./ Schnädelbach, Herbert (2000) (Hrsg.): *Naturalismus*. Philosophische Beiträge, Frankfurt am Main: Suhrkamp.

Keller, Fox Evelyn (2005): "The Century Beyond the Gene", in: *Journal of Bioscience* 30 (1): 3-10.

Kellert, Stephen H./ Longino Helen E./ Waters, C. Kenneth (2006): "Introduction: The Pluralist Stance", in: Kellert, Stephen H./ Longino Helen E./ Waters, C. Kenneth (eds.): *Scientific Pluralism.* Minnesota Studies in the Philosophy of Science, Minneapolis/ London: University of Minnesota Press, vii-xxix.

Kennedy, S. G./ Wagner, A. J./ Conzen, S. D./ Jordan, J./ Bellacosa, A./ Tsichlis, P. N./ Hay, N. (1997): "The PI 3-kinase/ Akt Signaling Pathway Delivers an Anti-Apoptotic Signal", in: *Genes & Development* 11: 701-713.

Kim, Jaegwon (1989): "The Myth of Non-Reductive Materialism", in: *Proceedings and Addresses of the American Philosophical Association* 63 (3): 31-47.

Kim, Jaegwon (1993): *Mind and Supervenience*, Cambridge: Cambridge University Press.

Kim, Jaegwon (1999): "Making Sense of Emergence," in: *Philosophical Studies* 95: 3-36.

Kim, Jaegwon (2005): *Physicalism, Or Something Near Enough*, Princeton: Princeton University Press.

Kim, Jaegwon (2006): *Philosophy of Mind.* Second Edition, Oxford: Westview.

Kim, Jaegwon (2008): "Reduction and Reductive Explanation: Is One Possible Without the Other?", in: Hohwy, Jakob/ Kallestrup, Jesper (eds.): *Being Reduced.* New Essays on Reduction, Explanation, and Causation, Oxford/ New York: Oxford University Press, 91-114.

Kincaid, Harold (1990): "Molecular Biology and the Unity of Science", in: *Philosophy of Science* 57 (4): 575-593.

Kitano, Hiroaki. (2002): "Systems Biology: A Brief Overview", in: *Science* 295: 1662-1664.

Kitcher, Philip (1981): "Explanatory Unification", in: *Philosophy of Science* 84: 507-531.

Kitcher, Philip (1984): "1953 an All That: A Tale of Two Sciences", in: *Philosophical Review* 93: 335-373.

Kitcher, Philip (1989): "Explanatory Unification and the Causal Structure of the World", in: Kitcher, Philip/ Salmon, Wesley C. (eds.): *Scientific Explanation*, Minneapolis: University of Minnesota Press, 410-506.

Kitcher, Philip (1993): "Function and Design", in: *Midwest Studies in Philosophy* 18 (1): 379-397.

Kitcher, Philip (1999a): "The Hegemony of Molecular Biology", in: *Biology and Philosophy* 14: 195-210.

Kitcher, Philip (1999b): "Unification as a Regulative Ideal", in: *Perspectives on Science* 7 (3): 337-348.

Kitcher, Philip (2011): “Philosophy Inside Out”, in: *Metaphilosophy* 42 (3): 248-260.

Kitcher, Philip/ Salmon, Wesley C. (1987): “Van Fraassen on Explanation”, in: *The Journal of Philosophy* 84 (6): 315-330.

Klein, Colin (2009): “Reduction Without Reductionism: A Defense of Nagel on Connectability”, in: *The Philosophical Quarterly* 59: 39-53.

Korfiatis, Kostas. J. / Stamou George. P. (1999): “Habitat Templets and the Changing Worldview of Ecology“, in: *Biology and Philosophy* 14: 375-393.

Kornblith, Hilary (2007): “Naturalism and Intuitions”, in: *Grazer Philosophische Studien* 74 (1): 27-49.

Krohs, Ulrich (2004): *Eine Theorie Biologischer Theorien.* Status und Gehalt von Funktionsaussagen und informationstheoretischen Modellen, Berlin: Springer.

Ladyman, James/ Ross, Don/ Spurrett, David/ Collier, John (2007): *Everything Must Go.* Metaphysics Naturalized, Oxford: Oxford University Press.

Lange, Marc (2000): *Natural Laws in Scientific Practice*. Oxford: Oxford University Press.

Laplace, Pierre Simon (1951): *A Philosophical Essay on Probabilities.* New York: Dover Publications.

Laubichler, Manfred D./ Wagner, Günter P. (2001): “How Molecular is Molecular Developmental Biology? A Reply to Alex Rosenberg’s ‘Reductionism Redux: Computing the Embryo’”, in: *Biology and Philosophy* 16: 53-68.

Levenstein, Susan (2009): “Against Reductionism”, in: *BMJ* 339: 709.

Levins, Richard (1970): “Complex Systems”, in: Waddington, C. H. (ed.): *Towards a Theoretical Biology*, Edinburgh: University Press, 73-88.

Levins, Richard/ Lewontin, Richard C. (1980): “Dialectics and Reductionism in Ecology”, in: *Synthese* 43: 47-78.

Levins, Richard/ Lewontin, Richard C. (1985): *The Dialectical Biologist,* Cambridge: Harvard University Press.

Levins, Richard/ Lewontin, Richard C. (2007): *Biology Under the Influence.* Dialectical Essays on Ecology, Agriculture, and Health, New York: Monthly Review Press.

Lewis, David (1983): “New Work for a Theory of Universals”, in: *Australasian Journal of Philosophy* 61 (4): 343–377.

Lewis, David (1986): “Events”, in: Lewis, David (ed.): *Philosophical Papers.* Vol. II, Oxford: Oxford University Press, 241-269.

Lewis, David (1994): “Reduction in Mind”, in: Guttenplan, Samuel (Ed.): *A Companion to Philosophy of Mind*, Oxford: Blackwell, 412-431.

Lidicker, William Z. (1988): “The Synergistic Effects of Reductionist and Holistic Approaches in Animal Ecology”, in: *OIKOS* 53: 278-281.

Lloyd, Elisabeth (1988): *The Structure and Confirmation of Evolutionary Theory*, New York: Greenwood Press.

Loehle, Craig (1988): “Philosophical Tools: Potential Contributions to Ecology”, in: *OIKOS* 51 (1): 97-119.

Love, Alan C. (2008a): “From Philosophy to Science (to Natural Philosophy): Evolutionary Developmental Perspectives”, in: *The Quarterly Review of Biology* 83 (1): 65-76.

Love, Alan C. (2008b): “Review Symposium: More Worry and Less Love”, in: *Metascience* 17: 1–26.

Love, Alan C. (2012): “Formal and Material Theories in Philosophy of Science: A Methodological Interpretation”, in: de Regt, Henk K./ Hartmann, Stephan/ Okasha, Samir (eds.): *EPSA Philosophy of Science: Amsterdam 2009,* Dordrecht: Springer, 175-185.

Love, Alan C./ Hüttemann, Andreas (2011): “Comparing Part-Whole Reductive Explanations in Biology and Physics” in: Dieks, Dennis/ Gonzalez, Wenceslao J./ Hartmann, Stephan/ Uebel, Thomas/ Weber, Marcel (eds.): *Explanation, Prediction, and Confirmation.* New Trends and Old Ones Reconsidered, Berlin: Springer, 183-202.

Machamer, Peter/ Darden, Lindley/ Craver, Carl F. (2000): “Thinking About Mechanisms”, in: *Philosophy of Science* 67: 1-25.

Machmer, Peter (2004): “Activities and Causation: The Metaphysics and Epistemology of Mechanisms”, in: *International Studies in the Philosophy of Science* 18 (1): 27-39.

Macilwain, Colin (2009): “Watching Science at Work”, in: *Nature* 462: 840-842.

Mackie, John L. (1974): *The Cement of the Universe*. A Study of Causation, Oxford: Clarendon Press.

Malaterre, Christophe (2007): “Organicism and Reductionism in Cancer Research: Towards a Systemic Approach”, in: *International Studies in the Philosophy of Science* 21 (1): 57-73.

Marcum, James A. (2005): “Metaphysical Presuppositions and Scientific Practices: Reductionism and Organicism in Cancer Research”, in: *International Studies in the Philosophy of Science* 19 (1): 31-45.

Mayr, Ernst (1982): *The Growth of Biological Thought*. Diversity, Evolution, and Inheritance, Cambridge: Belknap Press.

Mayr, Ernst (1988): “The Limits of Reductionism”, in: *Nature* 331: 475-476.

Mayr, Ernst (2004): *What Makes Biology Unique?* Considerations on the Autonomy of a Scientific Discipline, Cambridge: Cambridge University Press.

Mazzocchi, Fulvio (2008): “Complexity in Biology. Exceeding the Limits of Reductionism and Determinism Using Complexity Theory”, in: *EMBO reports* 9: 10-14.

McClelland, Erin E./ Bernhardt, Paul/ Casadevall, Arturo (2005): “Coping with Multiple Virulence Factors: Which Is Most Important?”, in: *PLoS Pathogens* 1: 287-288.

McIntosh, Robert P. (1987): “Pluralism in Ecology”, in: *Annual Review of Ecology and Systematics* 18: 321-341.

McLaughlin, Brian/ Bennett, Karen (2011): “Supervenience”, in: Zalta, Edward N. (ed.): *The Stanford Encyclopedia of Philosophy* (Winter 2011 Edition), URL = <http://plato.stanford.edu/archives/win2011/entries/supervenience/>.

McLaughlin, Peter (2001): *What Functions Explain:* Functional Explanation and Self-Reproducing Systems, Cambridge: Cambridge University Press.

Mikkelson, Gregory. M. (2004): “Biological Diversity, Ecological Stability, and Downward Causation”, in: Oksanen, M./ Pietarinen, J. (eds.): *Philosophy and Biodiversity*, Cambridge: Cambridge University Press, 119-129.

Mikkelson, Gregory. M. (forthcoming): “Part-Whole Relationships and the Unity of Ecology”, in: Skipper, R. A. Jr./ Allen, C./ Ankeny, R. /Craver, C. F./ Darden, L. / Mikkelson, G. M./ Richardson, R. C. (eds.): *Philosophy Across the Life Sciences*, MIT Press.

Mitchell, Sandra D. (1997): “Pragmatic Laws”, in: *Philosophy of Science* 64: 468-479.

Mitchell, Sandra D. (2003): *Biological Complexity and Integrative Pluralism*. Cambridge: Cambridge University Press.

Mitchell, Sandra D. (2009): *Unsimple Truths.* Science, Complexity, and Policy, Chicago/ London: University of Chicago Press.

Mitchell, Sandra D./ Dietrich, Michael (2006): “Integration without Unification: An Argument for Pluralism in the Biological Sciences”, in: *American Naturalist* 168: 73-79.

Morgan, Mary/ Morrison, Margaret (1999): *Models as Mediators.* Perspectives on Natural and Social Science, Cambridge: Cambridge University Press.

Mühlhölzer, Felix (2005): “Naturalismus und Lebenswelt – Plädoyer für eine rein deskriptive Wissenschaftstheorie“, in: Gesang, Bernward (Hrsg.): *Deskriptive oder normative Wissenschaftstheorie?* Frankfurt: ontos, 49-73.

Nagel, Ernest (1952): “Wholes, Sums, and Organic Unities”, in: *Philosophical Studies* 3 (2): 17-32.

Nagel, Ernest (1961): *The Structure of Science.* Problems in the Logic of Scientific Explanation, London: Routledge.

Nickles, Thomas (1973): “Two Concepts of Intertheoretic Reduction”, in: *The Journal of Philosophy* 70 (7): 181-201.

Noble, Denis (2002): “Modeling the Heart – from Genes to Cells to the Whole Organ”, in: *Sciene* 295: 1678-1682.

Norton, John (2003): “A Material Theory of Induction”, in: *Philosophy of Science* 70 (4): 647-670.

Nurse, Paul (1997): “The Ends of Understanding”, in: *Nature* 387: 657.

O’Malley, Maureen A./ Dupré, John (2005): “Fundamental issues in systems biology”, in: *BioEssays* 27: 1270-1276.

Oppenheim, Paul/ Putnam, Hilary (1958): “Unity of Science as a Working Hypothesis”, in: Feigl, H./ Scriven, M./ Maxwell, G. (eds.): *Concepts, Theories and the Mind-Body Problem*. Minnesota Studies in the Philosophy of Science II, Minneapolis: University of Minnesota Press, 3-36.

Papineau, David (2009): “Naturalism”, in: Zalta, Edward N. (ed.): *The Stanford Encyclopedia of Philosophy* (Spring 2009 Edition), URL = <http://plato.stanford.edu/archives/spr2009/entries/naturalism/>.

Pereboom, Derk (2002): “Robust Nonreductive Materialism”, in: *Journal of Philosophy* 99: 499-531.

Pereboom, Derk/ Kornblith, Hilary (1990): “The Metaphysics of Irreducibility”, in: *Philosophical Studies* 63: 125-145.

Perini, Laura (2005): “Explanation in Two Dimensions: Diagrams and Biological Explanation”, in: *Biology and Philosophy* 20: 257-269.

Perini, Laura (forthcoming): “Diagrams in Biology”, in: *Knowledge Engineering Review*.

Place, Ullin T. (1956): “Is Consciousness a Brain Process?”, in: *British Journal of Psychology* 47 (1): 44-50.

Plantinga, Alvin (1996): “Methodological Naturalism?”, in: Van der Meer, J. (ed.): *Facets of Faith and Science*. Lanham: University Press of America.

Popper, Karl R. (1974): “Scientific Reduction and the Essential Incompleteness of all Science”, in: Ayala, Francisco J./ Dobzhansky, Theodosius G. (eds.): *Studies in the Philosophy of Biology.* Reduction and Related Problems, Berkeley: University of California Press, 259-282.

Potochnik, Angelika (2009): “Levels of Explanation Reconceived”, in: *Philosophy of Science* 77: 59-72.

Powell, Kendall. (2004): “All Systems Go”, in: *The Journal of Cell Biology* 165: 299-303.

Pradeu, Thomas (2010): “What is an organism? An immunological answer”, in: *History and Philosophy of the Life Sciences* 32: 247-268.

Putnam, Hilary (1975): *Mind, Language, and Reality*. Philosophical Papers. Vol. 2, New York: Cambridge University Press.

Railton, Peter (1981): “Probability, Explanation, and Information”, in: *Synthese* 48: 233-256.

Reece, Jane B./ Urry, Lisa A./ Cain, Michael L./ Wasserman, Steven A./ Minorsky, Peter V./ Jackson, Robert B. (2011): *Campbell Biology*. 9th Edition, Boston: Pearson.

Reichenbach, Hans (1938): “On Probability and Induction”, in: *Philosophy of Science* 5 (1): 21-45.

Rosenberg, Alexander (1985): *The Structure of Biological Science*. Cambridge: Cambridge University Press.

Rosenberg, Alexander (1994): *Instrumental Biology or the Disunity of Science*. Chicago: University of Chicago Press.

Rosenberg, Alexander (1997): “Reductionism Redux: Computing the Embryo”, in: *Biology and Philosophy* 12: 445-470.

Rosenberg, Alexander (2001): “How Is Biological Explanation Possible?”, in: *British Journal for Philosophy of Science* 52, 735-760.

Rosenberg, Alexander (2006): *Darwinian Reductionism*. Or, How to Stop Worrying and Love Molecular Biology, Cambridge: University of Chicago Press.

Rosenberg, Alexander/ Kaplan, David M. (2005): “How to Reconcile Physicalism and Antireductionism about Biology”, in: *Philosophy of Science* 72: 43-68.

Rosenberg, Alexander/ McShea, Daniel W. (2008): *Philosophy of Biology.* A Contemporary Introduction, New York: Routledge.

Roukos, Dimitrios H. (2011): “Networks medicine: from reductionism to evidence of complex dynamic biomolecular interactions”, in: *Pharmacogenomics* 12 (5): 695-698.

Ruse, Michael (1976): “Reduction in Genetics “, in: *PSA* 1974: 633-651.

Salmon, Wesley (1993): “Scientific Explanation and the Causal structure of the World”, in: Hillel-Ruben, David (ed.): *Explanation*, Oxford: Oxford University Press, 78-112.

Salmon, Wesley C. (1971): “Statistical Explanation”, in: Salmon, Wesley C. (ed.): *Statistical Explanation and Statistical Relevance*, Pittsburgh: University of Pittsburgh Press, 29-87.

Salmon, Wesley C. (1977): “A *Third Dogma of Empiricism*”, in: Butts, R./ Hintikka J. (eds.): *Basic Problems in Methodology and Linguistics*, Dordrecht: Reidel, 149-166.

Salmon, Wesley C. (1984a): *Scientific Explanation and the Causal Structure of the World*, Princeton: Princeton University Press.

Salmon, Wesley C. (1984b): “Scientific Explanation: Three Basic Conceptions”, in: *PSA* 1984 (2): 293-305.

Salmon, Wesley C. (1989): “Four Decades of Scientific Explanation”, in: Kitcher, Philip/ Salmon, Wesley C. (eds.): *Scientific Explanation*. Minneapolis: University of Minnesota Press, 3-219.

Salmon, Wesley C. (1992): “Scientific Explanation”, in: Salmon, M. H./ Earman, J./ Glymour, C./ Lennox, J. G./ Machamer, Peter/ McGuire, J. E. /Norton, J. D./ Salmon, W. C./ Schaffner, K. F. (eds.): *Introduction to the Philosophy of Science*. Indianapolis/ Cambridge: Hackett Publishing, 7-41.

Salmon, Wesley C. (1993): “Scientific Explanation and the Causal Structure of the World”, in: Hillel-Ruben, David (ed.): *Explanation.* Oxford Readings in Philosophy, Oxford: Oxford University Press, 78-112.

Salmon, Wesley C. (1994): “Causality Without Counterfactuals”, in: *Philosophy of Science* 61, 297-312.

Salmon, Wesley C. (1997): “Causality and Explanation: A Reply to Two Critiques”, in: *Philosophy of Science* 64, 461-477.

Salmon, Wesley C. (1998): *Causality and Explanation*, Oxford: Oxford University Press.

Sarkar, Sahotra (1992): “Models of Reduction and Categories of Reductionism”, in: *Synthese* 91: 167-194.

Sarkar, Sahotra (1998): *Genetics and Reductionism*, Cambridge: Cambridge University Press.

Sarkar, Sahotra (2005): *Molecular Models of Life.* Philosophical Papers on Molecular Biology, Cambridge/ London: MIT Press.

Sarkar, Sahotra (2008): “Reduction”, in: Psillos, Stathis/ Curd, Martin (eds.): *The Routledge Companion to Philosophy of Science*, London/ New York: Routledge, 425-434.

Sarkar, Sahotra (2009): “Ecology“, in: Zalta, Edward N. (ed.): *Stanford Encyclopedia of Philosophy (Spring 2009 Edition)*, URL = <http://plato.stanford.edu/archives/spr2009/entries/ecology/>.

Schaffner, Kenneth F. (1967): “Approaches to Reduction”, in: *Philosophy of Science* 34: 137-147.

Schaffner, Kenneth F. (1969): “The Watson-Crick Model and Reductionism”, in: *British Journal for the Philosophy of Science* 20: 325-348.

Schaffner, Kenneth F. (1974a): “The Peripherality of Reductionism in the Development of Molecular Biology”, in: *Journal of the History of Biology* 7 (1): 111-139.

Schaffner, Kenneth F. (1974b): “Reductionism in Biology: Prospects and Problems”, in: *PSA* 1974: 613-632.

Schaffner, Kenneth F. (1993): *Discovery and Explanation in Biology and Medicine*, Chicago/ London: University of Chicago Press.

Schaffner, Kenneth F. (2006): “Reduction: The Cheshire Cat Problem and a Return to the Roots”, in: *Synthese* 151: 377-402.

Schoener, Thomas W. (1986): “Mechanistic Approaches to Community Ecology: A New Reductionism?”, in: *American Zoologist* 26: 81-106.

Schurz, Gerhard (2005): “Rationale Rekonstruktion: die Methode der Wissenschaftstheorie“, in: Gesang, Bernward (Hrsg.): *Deskriptive oder normative Wissenschaftstheorie?* Frankfurt: ontos, 123-144.

Scriven, Michael (1959): “Truisms as the Grounds of Historical Explanations”, in: Gardiner, P. (ed.): *The Nature of Historical Explanation*, Oxford: Oxford University Press, 443-475.

Service, Robert F. (1999): “Exploring the Systems Life”, in: *Science* 284: 80-83.

Shimamura, Haruko/ Terada, Yoshio/ Okado, Tomokazu/ Tanaka, Hiroyuk/ Inoshita, Seiji/ Sasaki, Sei (2003): “The PI3-kinase-Akt Pathway Promotes Mesangial Cell Survival and Inhibits Apoptosis in Vitro Via NF-kappa B and Bad”, in: *Journal of American Society of Nephrology* 14 (6): 1427-1434.

Simon, Herbert A. (1962): "The Architecture of Complexity", in: *Proceedings of the American Philosophical Society* 106 (6): 467-482.

Simon, Herbert A. (1973): "The Organization of Complex Systems", in: Pattee, Howard H. (ed.): *Hierarchy Theory: The Challenge of Complex Systems*, New York: Braziller.

Simons, Peter (1987): *Parts*. A Study in Ontology, Oxford: Clarendon.

Skipper, Robert A./ Millstein, Roberta L. (2005): "Thinking about evolutionary mechanisms: natural selection", in: *Studies in the History and Philosophy of Biological and Biomedical Sciences* 36: 327-347.

Sloep, Peter/ Van der Steen, Wim (1987): "The Nature of Evolutionary Theory: The Semantic Challenge", in: *Biology and Philosophy* 2: 1-15.

Smart, J. J. C. (1959): "Sensations and Brain Processes", in: *Philosophical Review* 68: 141-56.

Smith, Barry/ Varzi, Achille C. (2000): "Fiat and Bona-Fide Boundaries", in: *Philosophy and Phenomenological Research* 60 (2): 401-420.

Sober, Elliott (1999): "The Multiple Realizability Argument Against Reductionism", in: P*hilosophy of Science* 66: 542-564.

Sober, Elliott (2008): *Evidence and Evolution.* The Logic behind the Science, Cambridge: Cambridge University Press.

Sonnenschein, Carlos/ Soto, Ana M. (2000): "Somatic Mutation Theory of Carcinogenesis: Why It Should be Dropped and Replaced", in: *Molecular Carcinogenesis* 29: 205-211.

Sorger, Peter K. (2005): "A Reductionist's Systems Biology", in: *Current Opinion in Cell Biology* 17: 9-11.

Soto, Ana M./ Rubin, Beverly S./ Sonnenschein, Carlos (2006): "Emergentism by Default: A View From the Bench", in: *Synthese* 151: 361-376.

Soto, Ana M./ Sonnenschein, Carlos (2005): "Emergentism as a Default: Cancer as a Problem of Tissue Organization", in: *Journal of Bioscience* 30 (1): 103-118.

Soto, Ana M./ Sonnenschein, Carlos (2009): "Interpreting Endocrine Disruption From an Integrative Biology Perspective", in: *Molecular and Cellular Endocrinology* 304: 3-7.

Soto, Ana M./ Sonnenschein, Carlos (2010): "Environmental Causes of Cancer: Endocrine Disruptors as Carcinogenes", in: *Nature Reviews* 6: 363-370.

Spencer, Matthew (1997): "The effects of habitat size and energy on food web structure: An individual-based cellular automata model", in: *Ecological Modelling* 94: 299-316.

Steel, Daniel P. (2004): "Can a reductionist be a pluralist?", in: *Biology an Philosophy* 19: 55-73.

Steel, Daniel P. (2008): *Across the Boundaries*. Extrapolation in Biology and Social Science, Oxford/ New York: Oxford University Press.

Sterelny, Kim/ Griffiths, Paul E. (1999): *Sex and Death.* An Introduction to Philosophy of Biology, Chicago: University of Chicago Press.

Stoljar, Daniel (2009): “Physicalism”, in: Zalta, Edward N. (ed.): *The Stanford Encyclopedia of Philosophy* (Fall 2009 Edition), URL = <http://plato.stanford.edu/archives/fall2009/entries/physicalism/>.

Stotz, Karola/ Griffiths, Paul E. (2004): “Representing Genes: Phase I 2002-2004: Testing Competing Philosophical Analyses of the Gene Concept in Contemporary Molecular Biology; Phase II 2003-2005: Conceptual Issue in the Dissemination and Reception of Genomics”, URL = <http://www.pitt.edu/~kstotz/genes/genes.html>.

Stotz, Karola/ Griffiths, Paul E. (2005): “Genes: Philosophical Analyses Put to the Test”, in: *History and Philosophy of the Life Sciences* 26: 5-28.

Strange, Kevin (2005): “The End of ‘Naive Reductionism’: Rise of Systems Biology or Renaissance of Physiology?”, in: *American Journal of Physiology - Cell Physiology* 288: 968-974.

Strevens, Michael (2006): “Scientific Explanation”, in: Borchert, Donald M. (ed.): *Encyclopedia of Philosophy*. 2nd Edition, Detroit: Macmillan.

Strevens, Michael (2008): *Depth*. An Account of Scientific Explanations, Cambridge/ London: Harvard University Press.

Suppe, Frederick (1977): *The Structure of Scientific Theories*. 2nd Edition, Urbana: University of Illinois Press.

Suppe, Frederick (1989): *The Semantic Conception of Theories and Scientific Realism*, University of Illinois Press: Chicago.

Suppe, Frederick (2000): “Understanding Scientific Theories: An Assessment of Developments, 1969-1998”, in: *Philosophy of Science* 67: 102-115.

Tabery, James G. (2004): “Synthesizing Activities and Interactions in the Concept of Mechanism”, in: *Philosophy of Science* 71: 1-15.

Thompson, Paul (1989): *The Structure of Biological Theories*, Albany: State University of New York Press.

Torres, Phillip J. (2008): “A Modified Conception of Mechanism“, in: *Erkenntnis* 71 (2): 233-251.

van Fraassen, Bas C. (1977): “The Pragmatics of Explanation”, in: *American Philosophical Quarterly* 14 (2): 143-150.

van Fraassen, Bas C. (1980): *The Scientific Image*, Oxford: Oxford University Press.

van Regenmortel, Marc H. V. (1998): “From Absolute to Exquisite Specificity. Reflection on the Fuzzy Nature of Species, Specificity and Antigenic Sites”, in: *Journal of Immunological Methods* 216: 37-48.

van Regenmortel, Marc H. V. (2004): “Reductionism and Complexity in Molecular Biology”, in: *EMBO reports* 5: 1016-1020.

van Regenmortel, Marc H. V./ Hull, David L. (eds.) (2002): *Reductionism in the Biomedical Sciences*, London: John Wiley & Sons.

Venter, J. Craig. et al. (2001): "The Sequence of the Human Genome", in: *Science* 291: 1304-1348.

Waters, C. Kenneth (1990): "Why the Antireductionist Consensus Won't Survive the Case of Classical Mendelian Genetics", in: *PSA* 1990 (1): 125-139.

Waters, C. Kenneth (1994): "Genes Made Molecular", in: *Philosophy of Science* 61 (2): 163-185.

Waters, C. Kenneth (2004): "What Concept Analysis in Philosophy of Science Should Be (and Why Competing Philosophical Analyses of Gene Concepts Cannot Be Tested by Polling Scientists)", in: *History and Philosophy of the Life Sciences* 26 (1): 29-58.

Waters, C. Kenneth (2007): "Causes That Make a Difference", in: The *Journal of Philosophy* 104 (11): 551-579.

Waters, C. Kenneth (2008): "Beyond Theoretical Reduction and Layer-Cake Antireduction: How DNA Retooled Genetics and Transformed Biological Practice", in: Ruse, Michael (ed.): *The Oxford Handbook of the Philosophy of Biology*, Oxford: Oxford University Press, 238-262.

Watson, James D./ Crick, Francis H. C. (1953): "A Structure for Deoxyribose Nucleic Acid", in: *Nature* 171, 737-738.

Weber, Marcel (1998): *Die Architektur der Synthese.* Entstehung und Philosophie der modernen Evolutionstheorie, Berlin/ New York: de Gruyter.

Weber, Marcel (2005): *Philosophy of Experimental Biology,* Cambridge/ New York: Cambridge University Press.

Weber, Marcel (2008): "Critical Notice: Darwinian Reductionism", in: *Biology and Philosophy* 23 (1): 143-152.

Weinberg, Robert A. (1998): *One Renegade Cell*. New York: Basic Books.

Wetzel, Linda (2011): "Types and Tokens", in: Zalta, Edward N. (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), URL = <http://plato.stanford.edu/archives/spr2011/entries/types-tokens/>.

Wheeler, William Morton (1920): "Termitodoxa, or Biology and Society", in: *The Scientific Monthly* 10 (2): 113-124.

Wilson, David Sloan (1975): "A Theory of Group Selection", in: *Proceedings of the National Academy of Sciences USA* 72: 143-146.

Wilson, David Sloan (1988): "Holism and reductionism in evolutionary ecology", in: *OIKOS* 53 (2): 269-273.

Wilson, Jessica M. (2010): "Non-reductive Physicalism and Degrees of Freedom", in: *British Journal for Philosophy of Science* 61 (2): 279-311.

Wilson, Robert A. (2005): *Genes and the Agents of Life: The Individual in the Fragile Sciences*, Cambridge: Cambridge University Press.

Wilson, Robert A. (2008): “The Biological Notion of Individual”, in: Zalta, Edward N.(ed.): *Stanford Encyclopedia of Philosophy*. (Fall 2008 Edition), URL = <http://plato.stanford.edu/archives/fall2008/entries/biology-individual/>.

Wimsatt, William C. (1974): “Complexity and Organization“, in: *PSA* 1972: 67-86.

Wimsatt, William C. (1976a): “Reductive Explanation: A Functional Account”, in: *PSA* 1974: 671-710.

Wimsatt, William C. (1976b): “Reductionism, Levels of Organization, and the Mind-Body Problem”, in: Globus, Gordon G. (ed.): *Consciousness and the Brain*, New York/ London: Plenum Press, 205-267.

Wimsatt, William C. (1980): “Reductionistic Research Strategies and Their Biases in the Units of Selection Controversy”, in: Nickles, Thomas (ed.): *Scientific Discovery: Case Studies*, Dordrecht: Reidel, 213-259.

Wimsatt, William C. (1981): “Robustness, Reliability and Overdetermination”, in: Brewer, M./ Collins, B. (eds.): *Scientific Inquiry and the Social Sciences*, San Francisco: Jossey-Bass Publishers, 124-163.

Wimsatt, William C. (1986): “Forms of Aggregativity”, in: Donagan, Alan/ Perovich, Anthony N. Jr./ Wedin, Michael V. (eds.): *Human Nature and Natural Knowledge.* Essays Presented to Marjorie Grene on the Occasion of Her Seventy-Fifth Birthday, Boston: Reidel, 259-291.

Wimsatt, William C. (1997): “Aggregativity: Reductive Heuristics for Finding Emergence”, in: Philosophy of Science 64: 372-384.

Wimsatt, William C. (2006a): “Reductionism and its Heuristics: Making Methodological Reductionism Honest”, in: *Synthese* 151 (3): 445-475.

Wimsatt, William C. (2006b): “Aggregate, composed, and evolved systems: Reductionistic heuristics as means to more holistic theories”, in: *Biology and Philosophy* 21 (5): 667-702.

Wimsatt, William C. (2007): *Re-Engineering Philosophy for Limited Beings.* Piecewise Approximations to Reality, Cambridge: Harvard University Press.

Wimsatt, William C./ Sarkar, Sahotra (2006): “Reductionism”, in: Sarkar, Sahotra/ Pfeifer, Jessica (eds.): *The Philosophy of Science: An Encyclopedia*, New York: Routledge, 696-702.

Winther, Rasmus Grønfeld (2006): “Parts and theories in compositional biology”, in: *Biology and Philosophy* 21: 471-199.

Winther, Rasmus Grønfeld (2011): “Part-whole science”, in: *Synthese* 178: 397-427.

Woodward, James (2003): *Making Things Happen*. A Theory of Causal Explanation, Oxford: Oxford University Press.

Woodward, James (2010): “Causation in Biology: Stability, Specificity, and the Choice of Levels of Explanation”, in: *Biology and Philosophy* 25 (3): 287-318.

Woodward, James (2011): “Scientific Explanation”, in: Zalta, Edward N. (ed.): *The Stanford Encyclopedia of Philosophy* (Winter 2011 Edition), URL = <http://plato.stanford.edu/archives/win2011/entries/scientific-explanation/>.

Yashiro, Kenta/ Shiratori, Hidetaka/ Hamada, Hiroshi (2007): “Haemodynamics determined by a genetic programme govern asymmetric development of the aortic arch”, in: *Nature* 450: 285-288.

Yoo, Julie (2008): “New Hope for Non-Reductive Pysicalism”, in: Hieke, Alexander/ Leitget, Hannes (eds.): *Papers of the 31st International Wittgenstein Symposium: Reduction and Elimination in Philosophy and the Sciences.* Vol. 16, Kirchberg am Wechsel: Eigner, 408-411.