History, Philosophy and Theory of the Life Sciences

Marie I. Kaiser

Reductive Explanation in the Biological Sciences

Springer

History, Philosophy and Theory of the Life Sciences

Volume 16

Editors
Charles T. Wolfe, Ghent University, Belgium
Philippe Huneman, IHPST (CNRS/Université Paris I Panthéon-Sorbonne), France
Thomas A.C. Reydon, Leibniz Universität Hannover, Germany

Editorial Board
Marshall Abrams (University of Alabama at Birmingham)
Andre Ariew (Missouri)
Minus van Baalen (UPMC, Paris)
Domenico Bertoloni Meli (Indiana)
Richard Burian (Virginia Tech)
Pietro Corsi (EHESS, Paris)
François Duchesneau (Université de Montréal)
John Dupré (Exeter)
Paul Farber (Oregon State)
Lisa Gannett (Saint Mary's University, Halifax)
Andy Gardner (Oxford)
Paul Griffiths (Sydney)
Jean Gayon (IHPST, Paris)
Guido Giglioni (Warburg Institute, London)
Thomas Heams (INRA, AgroParisTech, Paris)
James Lennox (Pittsburgh)
Annick Lesne (CNRS, UPMC, Paris)
Tim Lewens (Cambridge)
Edouard Machery (Pittsburgh)
Alexandre Métraux (Archives Poincaré, Nancy)
Hans Metz (Leiden)
Roberta Millstein (Davis)
Staffan Müller-Wille (Exeter)
Dominic Murphy (Sydney)
François Munoz (Université Montpellier 2)
Stuart Newman (New York Medical College)
Frederik Nijhout (Duke)
Samir Okasha (Bristol)
Susan Oyama (CUNY)
Kevin Padian (Berkeley)
David Queller (Washington University, St Louis)
Stéphane Schmitt (SPHERE, CNRS, Paris)
Phillip Sloan (Notre Dame)
Jacqueline Sullivan (Western University, London, ON)
Giuseppe Testa (IFOM-IEA, Milano)
J. Scott Turner (Syracuse)
Denis Walsh (Toronto)
Marcel Weber (Geneva)

More information about this series at http://www.springer.com/series/8916

Marie I. Kaiser

Reductive Explanation in the Biological Sciences

Springer

Marie I. Kaiser
Universität zu Köln
Köln, Germany

ISSN 2211-1948 ISSN 2211-1956 (electronic)
History, Philosophy and Theory of the Life Sciences
ISBN 978-3-319-25308-4 ISBN 978-3-319-25310-7 (eBook)
DOI 10.1007/978-3-319-25310-7

Library of Congress Control Number: 2015957992

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

For Nils

Acknowledgments

This book emerged from my doctoral dissertation "An Ontic Account of Explanatory Reduction in Biology," submitted to the Philosophy Faculty of the University of Cologne in July, 2012. My dissertation project was supervised by Andreas Hüttemann (Cologne) and Marcel Weber (Geneva), and I am very grateful to both of them for their continuous support and advice.

Since 2012, I have substantially rethought, developed, and complemented my work on reductive explanation in the biological sciences. During this process, I got helpful, though sometimes conflicting, inspirations from many sides. My German research group working on "Causation, Laws, Dispositions, and Explanation at the Intersection of Science and Metaphysics" encouraged me to reflect on the metaphysical assumptions that underlie reductionistic explanatory and investigative practices in the biological sciences. I would like to thank my colleagues in this group for their inspiring discussions and commitment over the last 6 years. The financial support that made this book possible was provided by the German Research Foundation (Deutsche Forschungsgemeinschaft; FOR-1063).

My aim to understand what reduction *in biological practice* is, rather than analyzing reduction in biology "from the armchair," arose and was reinforced during my visits at the University of Minnesota (USA). I am very grateful to Ken Waters and Alan Love for having acquainted me with and convinced me of the "practice turn" in philosophy of science (cf. Chap. 2, Sect. 1). Even though I still retain some of the monistic aspirations that are distinctive of metaphysicians of science and general philosophers of science (e.g., by proposing *one* account of reductive explanation that is supposed to hold for biology in general), the pluralistic philosopher of biology is always in my head, alerting me to take the variety of biological practice seriously.

My time at the University of Geneva (Switzerland) allowed me to intensify my studies of biological practice, with a special eye on the question of biological parthood. The "Lake Geneva Biological Interest Group (lgBIG)" assembles philosophers of biology, historians of science, and biologists and thereby created a very

stimulating surrounding for developing some of the ideas in this book further. I would like to express my gratitude to all my colleagues in Geneva for the wonderful time and for the intense and fruitful discussions we had.

I would also like to thank the editors of this book series and two anonymous referees for their helpful comments on earlier versions of the manuscript.

Cologne Marie I. Kaiser
August, 7th 2015

Contents

6 The Ontic Account of Explanatory Reduction 173

Chapter 1
Introduction

The goal of this book is to provide an understanding of an important element of contemporary biological research practice, namely of *explanatory reduction*, or more precisely, of *reductive explanations.*[1] My central question is: What makes an explanation in the biological sciences reductive and distinguishes it from non-reductive explanations?

The topic of reduction(ism) was and still is a much discussed issue in fields such as general philosophy of science, philosophy of mind, and 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 1970s. The general topic I am concerned with in this book is thus not new. But 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 twentieth 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 (cf. Kaiser 2012). Instead, I focus on answering a question that I think is prior to discussions about explanatory reduc-

---

[1] I assume that there are four types of reduction that one should keep apart: ontological reduction and three kinds of epistemic reduction, namely theory, methodological, and explanatory reduction (this difference is spelled out in Chap. 3).

M.I. Kaiser, *Reductive Explanation in the Biological Sciences*, History,
Philosophy and Theory of the Life Sciences 16,
DOI 10.1007/978-3-319-25310-7_1

tionism (see Chap. 3, Sect. 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.

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 and Dietrich 2006; Darden 2005; Bechtel and Richardson 2010). In my view, this is not the right strategy. Reductions – more precisely, reductive methods and reductive explanations – were and remain important 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 objects or 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 Chap. 6, Sect. 1.3). 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 Chap. 3). 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; see also Kaiser 2011) or call for a move "beyond reductionism" (Gallagher and 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 question of what makes an explanation reductive 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. 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 (at least if one is interested in in-practice kinds of explanatory reductionism; see Chap. 3, Sect. 1).

The general structure of my book is the following. Chapter 2 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 Chap. 3 is to introduce the previous debate about reduction(ism) in the philosophy of biology. But 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 Chap. 4 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 Chap. 5. 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 Chap. 6, develop my own account of explanatory reduction in biology. I will start with briefly specifying two concepts that occupy center stage in my account: the concept of a biological part (or of a part-whole relation) and the concept of levels of organization. On the basis of these conceptual clarifications I can then 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 of biological practice will be that reductive explanations in biology 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 biological object of interest, and describe the biological parts of this object only as parts in isolation.

This account deviates from previous discussions not only insofar as it does not discuss the truth or falsity of reductionism and is not concerned with reduction as a relation among theories. Moreover, it is innovative because it reveals three specific criteria of the reductivity of biological explanations, which have not been discussed in the literature in this way before. These three criteria point out very clearly what the reductive character of an explanation consists in, also because they trace the reductivity of explanations back to specific relations that exist in the world and that are represented in a certain way by reductive explanations. This is why I characterize my account as an *ontic* account of explanatory reduction.[2]

---

[2] Note that this understanding of the term 'ontic' deviates from how Craver and Strevens understand it. I do not claim, as they do, that explanations are entities in the world, rather than representations of entities in the world (more on this in Chap. 5, Sect. 2 and Chap. 6, Sect. 6).

Chapter 2
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)*

Contents

M.I. Kaiser, *Reductive Explanation in the Biological Sciences*, History, Philosophy and Theory of the Life Sciences 16,
DOI 10.1007/978-3-319-25310-7_2

The aim of this book is to understand an important element of contemporary biological research practice, namely reduction. The central questions of this book are: What is reduction in biology? (in other words, Which important characteristics of actual cases of reduction can be identified?), and Where do the strengths and limitations of reduction in biological practice lie? For reasons that I will reveal in detail later on (see Chap. 3), my analysis of reduction will focus on examining examples of reductive (and non-reductive) explanations that play an important role in biological practice. I use the term ‘biological (research) practice’ not in opposition to biological theory or theorizing, but rather in its broadest meaning. Biological practice comprises everything a biologist is typically engaged in when conducting research: narrowly “practical” elements (e.g., applying techniques, instruments, or investigative strategies), but also elements that may be characterized as more “theoretical” (e.g., building theories, modeling, or developing explanations).[1]

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 make these preliminary assumptions explicit by specifying how I proceed in developing my account of reduction, which methodology I use and which goals I pursue. Questions like this, that is, questions about how to do philosophy of science are located on a *meta-level* of philosophical analysis. On this meta-level 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[2] is. Meta-philosophical questions that are relevant in the context of reduction are for instance: Is the goal of a philosophical 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 a philosophical theory about 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 this book since it explicates and justifies the philosophical methodology by which I develop my account of explanatory reduction.

---

[1] One might also characterize scientific practice in terms of scientific activities, such as classifying, modelling, hypothesizing, idealizing, observing, and so on (Chang 2011).

[2] In the following I will use ‘account’ synonymously with ‘theory’ or ‘model’ in order to refer to a bunch of philosophical theses, for instance, about a certain element or feature of a particular scientific field or about science in general.

Although there is a long-standing debate about how philosophy in general is to be carried out (called meta-philosophy),[3] the specific question of how to do philosophy of science is rarely extensively debated or pursued in its own right.[4] 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. Hence, only a few of the meta-philosophical discussions in philosophy of science play more than a propaedeutic role. This is, however, 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. William Wimsatt, for instance, devotes an entire 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 scientific 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 strategy by which I develop 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) are legitimate philosophical projects as well (more on this in Sect. 2).[5]

---

[3] 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?

[4] This is particularly true with regard to the question of how to develop an account of reduction in biology.

[5] If not indicated otherwise section references refer to sections in the current chapter.

In order to give this 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* questions in the meta-philosophy of science that have not been satisfactorily and systematically explored so far. The second goal of this chapter is thus to contribute to filling this gap by introducing useful distinctions and clarifying relevant relationships. In so doing I hope to make 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 (Sect. 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 (Sect. 1.1). I clarify in which sense descriptive accounts in philosophy of science can be called naturalistic (Sect. 1.2) and why the best way to develop them is to conduct a philosophical analysis in a bottom-up fashion (Sect. 1.3). I conclude by specifying what the empirical data are that a descriptive account in philosophy of science needs to capture (Sect. 1.4).

In Sect. 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 spectrum of different kinds of philosophical projects. I argue that the end points of this spectrum (purely normative and purely descriptive projects) are empty and that the middle ground between descriptive-critical and normative projects is filled by projects that claim to be descriptive and normative at once.

In Sect. 3 I analyze one of these different kinds of projects, namely descriptive projects, in more detail. I argue that descriptive projects in philosophy of science cannot be *purely* descriptive since philosophers must take up a *critical* (and in a certain sense normative) stance on the empirical data scientific practice provides them with. 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 Sect. 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 (Sect. 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" (Sect. 4.2). Finally, I discuss the objection that this criterion is not pluralistic enough (Sect. 4.3).

In Sect. 5 I pick up on the issue of normativity again. The central task of this section is to clarify what it means when 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 as good (Craver 2007a). In order to clarify the notion of normativity that is involved here I first specify the concept of epistemic norms in science (Sect. 5.1). Then I distinguish three dimensions or kinds of normativity that can be observed in philosophy of science: methodological normativity, which is typical for projects that claim to be descriptive (or naturalistic), normativity in the strict sense, and norm-normativity, which can be divided into the critical description of norms and the independent justification of norms. In Sect. 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 in which way my own analysis of reductive explanations in biology is normative (Sect. 5.4).

In the last section of this chapter (Sect. 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

The aim of this book is to understand what reduction in actual biological practice is and where the strengths and limits of performing reductions lie. To pursue this goal presupposes *taking 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*). Hence, my project is part of philosophy of science that has undergone a "practice turn" (Soler et al. 2014) and that is also referred to as "philosophy of science in practice" (Ankeny et al. 2011). By identifying "understanding biological practice" as the aim of a philosophical account of reduction one thus accepts the following criterion of adequacy (which is further refined in Sect. 3):

Criterion of Descriptive Adequacy
An account of reduction in biology is adequate only if

(1) it captures cases of reduction that occur in current biological research practice.[6]

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).[7] 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 Sect. 3, I further specify what it means for a philosophical account to be descriptively adequate.

1.2 Methodological Naturalism

Instead of speaking about the descriptive character of their account many philosophers of science 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. Yet, it seems to me that the kind of naturalism that is relevant in this context of pursuing philosophy of science in a descriptive manner is what is referred to as *methodological naturalism*. According to David

---

[6] This also holds for other accounts in the philosophy of science.

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

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 and Schnädelbach 2000; Kornblith 2007), the general statement about philosophy remains largely the same: a methodological naturalist claims that with respect to their methods (and aims) there exists no principled difference between philosophy and the natural sciences. In other words, 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 critical discussion. 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 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. I thus 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.[8]

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. 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.[9] The difference is that in the natural sciences these empirical data are data about the natural world itself,[10] 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

---

[8] 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.

[9] We could thus also speak about a certain kind of *empirical adequacy* instead of descriptive adequacy.

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

causal inferences scientists make).[11] 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.

Still, 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 particular, two questions need to be answered in the remaining part of this chapter: one concerns the process of description, the other the phenomena to be described. First, how does theory building in philosophy of science (e.g., the construction of an account of reduction) 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 scientists, or philosopher's observations of 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

In Sect. 1.1 I argued that 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 and then moves up toward general philosophical claims. 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. 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

---

[11] 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).

philosophical thesis in a certain scientific field. When philosophy of biology emerged as a distinct discipline in the 1960s and 1970s 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 one's 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.[12]

1.4 The Data to Be Captured

Let us now turn toward the second of the two questions presented at the end of Sect. 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" 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., the concept of explanation, progress, model, complexity, law, experiment, mechanism), including concepts that are used as technical terms in the sciences itself (with respect to biology, e.g., the concept of 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 and McShea 2008, 4).

This, however, still leaves open the question of which empirical facts are to form the basis upon which the meaning of scientific concepts is 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. 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

---

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

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).[13] But 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 (cf. Einstein 1933, 5). James Woodward, for instance, emphasizes that his interventionist theory of causation goes beyond being a "mere" 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). 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; 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.[14]

2 Descriptive vs. Normative Projects in Philosophy of Science

An opponent of the 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*

---

[13] 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.

[14] 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 and 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".

of how science *should* proceed or, in the long run, *will* proceed (when seeking reductions)? He could insist that this philosophical ideal can be justified regardless of whether it is in fact realized in contemporary biological practice. And he could point to some important predecessors who have already argued in favor of such ideals of reduction. Kenneth F. Schaffner, for instance, 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 GRR model, required for theory reductions (Schaffner 1974a, 128). In spite of the peripherality of Schaffner's view of reduction, he clings to his GRR model as a *regulative ideal* that, even if it not yet is realized in biological practice, *should* guide the development of molecular biology (1993, 511).[15]

It seems to me that there are two options that must 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 (what this can mean is subject of Sect. 3). Let us call this the *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 much about what cases of reductions actually performed in biological practice look like. This is the *normative project*.

In my view 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. They might 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. 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 reduction ideally should be. What characterizes

---

[15] 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 largely a myth" (2006, 378). He claims that in the biological sciences we typically find "creeping reductions" (i.e. partial, multi-level reductive explanations) instead of "sweeping reductions" (2006, 397). But 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 extent by his GRR model. This point will be elaborated in Chap. 3, Sect. 4.

reduction *in practice* is ignored.[16] Figure 2.1 illustrates the difference between these two kinds of philosophical projects:

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). 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 (recall Sect. 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 and *then* the descriptive project is the only game in town. 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.

But the situation is more complicated than Fig. 2.1 suggests. Neither there is anything like a purely descriptive or a purely normative project in contemporary philosophy of science, nor can all projects pursued in philosophy of science be neatly divided into either the descriptive or the normative drawer. Let me elaborate on these points in turn. First, it may be that many (formal) philosophical theories about science developed in the heyday of Logical Empiricism were purely normative since they were developed without taking into account empirical information about actual scientific practice.[17] But nowadays in philosophy of

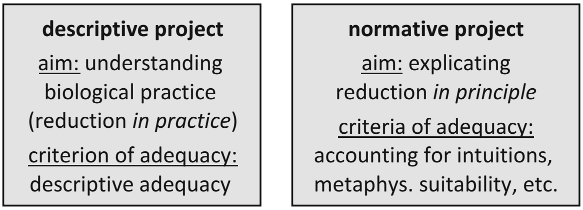

Fig. 2.1 Descriptive and normative projects in philosophy of science

---

[16] This difference between analyzing reduction *in practice* and developing an account of reduction *in principle* does *not* fully coincide with a difference that I will introduce in the next chapter, namely the difference between in practice and in principle claims in the reductionism debate. Even if you pursue a descriptive project and focus on the analysis of cases of reductions that are actually carried out *in practice*, it is still possible that you use this understanding of reduction to argue for *in principle* reductionism (e.g., for the thesis that, in principle, all biological explanations can be reduced to molecular explanations). But despite this possible combination, if you think it is important to reconstruct actual cases of reductions in order to understand what reduction is (focus on reduction *in practice*; descriptive project) it is likely that you will restrict your reductionist or antireductionist claims to cases of reductions that can be accomplished at present (*in practice* reductionism).

[17] 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).

science *purely normative* projects, in which facts about how science is actually undertaken are treated as completely irrelevant, are very rare. As already said, philosophers of science pursuing a normative project 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 projects of this kind are rarely pursued in a *purely* normative manner, that is, by ignoring any kind of empirical information about how science in fact works and why it is actually successful.[18] For instance, although Schaffner's GRR model of reduction clearly is a normative account (as it is not developed by reconstructing actual cases of reductions, but posits how reductions ideally should 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. The difference to descriptive projects, however, is that factual claims about the adaptation explanations that can be found in science do not play a noteworthy role in the development or justification of Brandon's normative account. It seems as if Schaffner as well as Brandon employs the examples for mere illustrative or refining purposes. To conclude, normative projects in philosophy of science are normative because they seek to develop normative claims about, for instance, what reduction in biology should be or how adaptation explanations should look like. Most normative projects, however, are not purely normative because factual claims, for example, about actual cases of reductions or adaptation explanations are included in the philosophical account, but they play no role in developing or justifying the normative claims.

In philosophy of science normative projects (even of the non-pure kind) are in the minority. This is true in particular with regard to the philosophy of the biological sciences. Philosophers of biology want to understand, for example, how the

---

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

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 – in other words, they pursue descriptive projects. But also descriptive projects are far from being carried out in a *purely descriptive* manner. Philosophers do not merely describe or passively depict scientific practice, but rather actively reconstruct it. Philosophical analysis requires taking up a selective and critical stance on the empirical information that can be gathered about scientific practice. In order to emphasize this critical stance and to indicate that there is no such thing as purely descriptive philosophy of science (even if purely descriptive projects were possible they would not be philosophical) I will speak of *descriptive-critical projects*. In the subsequent section I will elaborate on the critical character of descriptive philosophy of science and show that it amounts to making also normative claims concerning the selection of relevant empirical data. It is important to note that the dimension of normativity that is involved in descriptive-critical projects is different from the one that characterizes normative projects (see Sects. 3 and 5) – but still, descriptive projects are far from being purely descriptive.

Finally, the above picture of the two kinds of philosophical projects is too simplistic because it represents only two points of an entire spectrum of different kinds of philosophical projects. I have just argued that the two end points of this spectrum, purely descriptive and purely normative projects, can be disregarded – the former because they are not possible or at least not philosophical, the latter because they are very rare. But in the spectrum of projects in philosophy of science there is a middle ground between descriptive-critical and normative projects, which is filled with philosophical projects that claim to be descriptive and normative at once. This third kind of projects will be closely examined in Sect. 5. The variety of different projects is illustrated by Fig. 2.2.

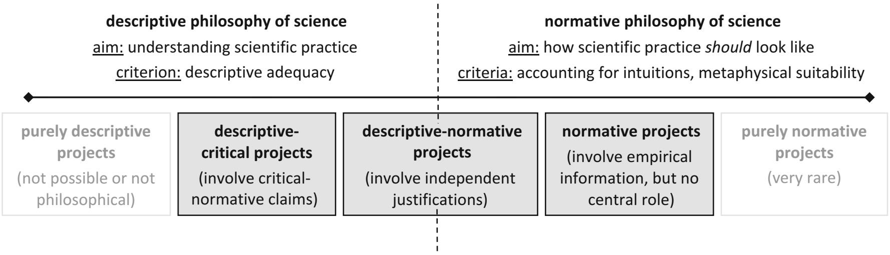

Fig. 2.2 A spectrum of different kinds of projects in philosophy of science

3 Why Pure Description Is Not Enough

Let us now come back to the first of the two questions identified at the end of Sect. 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? How can the criterion of descriptive adequacy, which was presented in Sect. 1.1, be further specified? For example, when does a philosophical theory succeed in 'capturing' real scientific cases? As I have shown in Sect. 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 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 describe all available cases and perhaps generalize 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*[19] or "explication" (Carnap 1950) rather than as a passive description of how science is done. 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 a trait of an organism to a function, what causation is, etc.) from scientific practice. The amount of available information about scientific practice is immense and overwhelming so that philosophers must find means to select those information that are relevant, for instance, because they represent paradigmatic, typical or important cases. This means that philosophers of science thus must take up a *critical stance* on the empirical data science provides them with. Moreover, frequently answers to philosophical questions cannot be easily found in scientific practice but require an act of interpretation and making explicit assumptions that are only implicit in scientific practice. Philosophers are also often confronted with a great deal of differences or even inconsistencies among scientific fields. Explanatory and investigative strategies vary, concepts are understood differently, and different background assumptions are made, etc. In order to develop a coherent, general account philosophers must take up a critical stance and sort out those empirical information about science that can be dismissed as false, misleading, or biased. Let us consider these different elements of the critical stance in more detail. At least three reasons can be given for why philosophy of science must go beyond the purely descriptive and exhibit a critical-normative character. I will explain them in the next three sections.

---

[19] The way I understand the term 'reconstruction' is similar to Hans Reichenbach's notion of a "rational reconstruction" (1938, 6; see also Schurz 2005) 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).

3.1 Focusing on Relevant Data

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 *relevant* examples, that is, examples that are "paradigmatic" (Wimsatt 2007, 27) and important. 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 Sect. 4). A philosopher should thus 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.[20]

3.2 Explicating Underlying Assumptions

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 theory about a certain element or feature of science must involve more than a mere description of scientific practice, namely a critical

---

[20] 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; Anderson 2005, 76 f; Janich 2005, 155f). I agree but it is important to recognize that this is just one respect in which philosophy of science can be normative and that there are other dimensions of normativity.

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 and 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 et al. 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 biological object or system by reference to its parts or about the method of analysis or decomposition. This reveals that what biologists regard as a reductive explanation and under which conditions they treat an explanation as reductive or as non-reductive often is only *implicit* in biological practice and needs to be unfolded by philosophers.

Against this background, the in Sect. 1.1 presented criterion of adequacy for an account of reduction needs to be refined:

Criterion of Descriptive Adequacy (refined)

If the aim is to understand biological practice, an account of reduction in biology is adequate only if

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

(b) it *explicates* the understanding of reduction that is (often only implicitly) present in biological practice.

3.3 Establishing Coherence

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). In order to provide a coherent and unified account, philosophers thus need to adopt a critical 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 (see Kaiser 2011). 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 Chap. 6, Sect. 4.2.1). 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-normative character.

3.4 Methodological Normativity

At this point one might ask: agreed, developing a (descriptive-naturalistic) philosophical theory about science always involves selecting relevant empirical information about science, making explicit assumptions that underlie scientific practice, and establishing coherence. But why should this critical-selective procedure be normative? Where exactly does normativity come into play? And does this turn descriptive projects into normative ones?

In my view the critical stance that descriptive projects in philosophy of science require clearly involves normative assumptions, for instance, about which examples should be conceived as paradigmatic, how the importance of an example for a certain field is assessed how scientific success and failure should be spelled out (assuming that importance is traced back to success/failure), and which statements of biologists should be conceived as correct or representative and thus should be included in the empirical basis of the philosophical account. In line with this, for instance 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).[21] But methodological normativity does not turn descriptive projects into normative projects (as characterized in Sect. 2) because the dimensions of normativity involved in both kinds of projects are different. Descriptive projects are normative in a different way that normative projects are normative. I call the kind or dimension of normativity involved in descriptive projects in philosophy of science *methodological normativity*. The normative claims made in descriptive projects merely concern methodological considerations, more precisely, they only concern the selection of those empirical information on the basis of which the philosophical account or theory about a certain subject X is developed. In descriptive projects not also normative claims *about the subject matter X itself* are made. These kind of normative claims are reserved for normative projects.

An example illustrates these two different dimensions of normativity. Schaffner's account of what theory reduction in the medical sciences is, his GRR model, is an example of a normative project in philosophy of science. It is normative because it contains normative claims about its subject matter (i.e., theory reduction). I call this *normativity in the strict sense*. For instance, the GRR model draws a picture of what theory reduction in the medical sciences should be; it does not represent how actual cases of theory reduction in this field in fact look like (which would be a factual claim, not a normative one). By contrast, a descriptive project that seeks to understand theory reduction in physics contains only factual claims about what the characteristics of actual cases of theory reductions in the physical sciences *are* (not what they should be). It is thus not normative in the strict sense that it makes normative claims about its subject matter. But it is normative in a different way since its development presupposes certain normative claims about how to select the relevant empirical information on the basis of which the account of theory reduction in physics is build (e.g., which examples of theory reductions that can be found in physics are paradigmatic and important and thus should be subject of philosophical analysis, which fields in physics should be regarded as epitome of successful science?). This methodological normativity is *inevitable* in descriptive philosophical projects – and in this sense there is no purely descriptive, normativity-free philosophy of science – but it can be hidden if the normative claims that concern the methodology of an account are not made explicit.

---

[21] 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.

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 pluralistically inclined philosophers could object: Why seek universality at all? They 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. But 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, two 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, the *scope of application* of the account is restricted to a specific field of science. For example, one might hold (as I do) that the question what a reductive explanation is needs to be answered differently with respect to, for instance, biology, physics, or the social 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. It thus 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 (I thus 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 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.[22] 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:

Criterion of Balancing Specificity Against Generality
If the aim is to understand biological practice, an account of reduction in biology 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 clearly indicates, 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

---

[22] 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.

and by not sweeping existing exceptions under the table. On the other hand it is the task of philosophers to abstract from some details of scientific research and to aspire after developing accounts that are coherent and general. Philosophy of science thus 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 in Sect. 4.1) 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. She 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*. 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).[23]

The thesis that philosophy of science should be pluralistic (or at least allow a plurality of philosophical accounts) is often motivated by the arguments for *pluralism in the sciences* itself (which is also the primary focus of the debate). Accordingly, the volume of Kellert et al. (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.[24] 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.

For the purposes of this chapter, however, it is not important whether or not scientific pluralism (as a metaphilosophical claim) can be defended or not. As should

---

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

[24] 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).

have become clear in the 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 Sect. 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 Sect. 1 I characterized the account of reduction I will develop in this book as descriptive, naturalistic, and bottom-up. In Sect. 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 actual cases of reduction that occur in biological practice, but rather posits what reduction ideally should be (normativity in the strict sense). In Sect. 3 I clarified that a descriptive project in philosophy of science must necessarily involve a critical stance and that this implies accepting certain normative assumptions about how to select relevant empirical information about science and about how to develop a coherent philosophical account. Section 4 specified the coherency-requirement.

The starting point of this section is the observation that several philosophers of science want to belong to the descriptive, naturalistic, bottom-up layer but, nevertheless, retain their normative aspirations (where normativity is not understood as mere methodological normativity). 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 Sect. 2. 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). These kind of descriptive-normative projects thus seem to lie in the middle between descriptive projects and normative projects (recall Sect. 2, Fig. 2.2). Jointly descriptive and normative projects seem to differ from descriptive projects in so far as they contain these additional normative claims. They differ from normative projects in so far as the normative claims they include are not mere philosophical ideals that are detached from real science. Rather, the normative claims are supposed to be "informed" by the results of the descriptive part of the project.

A major aim of this section is to analyze these projects that claim to be jointly descriptive and normative more closely, to uncover in which way they can be characterized as being normative, and to scrutinize what it means that the normative part of the project is "informed" by the descriptive part and whether this is a feasible relation. Since this will require of us to think more about issues of normativity and epistemic norms a second aim of this section is to continue the analysis in Sect. 3.4 and to specify different dimensions of normativity that can be found in philosophy of science.

If some philosophers hear the term 'normative philosophy of science' they associate this with debates about the influence of social or epistemic norms on science (e.g., on the development of scientific knowledge). The intuitive idea is that a philosophical theory about science is normative if its subject matter are social/epistemic norms and the role these norms play in scientific practice. I don't think that this idea is wrong, it is just too simplistic because the normativity of philosophy of science is multi-dimensional (more on this in Sect. 5.2). Still, when analyzing the normativity of philosophy of science one should also take into account (epistemic) norms in science. The reason why I discuss epistemic norms in this section is that many proponents of jointly descriptive and normative projects refer to epistemic norms (or virtues) to explain what the normative character of their account is. Craver, for instance, states that a major aim of his account of mechanistic explanation is to "provide a… set of norms by which [mechanistic] explanations should be assessed" (2007a, 111). Other philosophers of science explicitly demand that philosophical analyses of scientific concepts, such as the concept of mechanism or of a gene, must disclose the relation these concepts have to epistemic norms (Waters 2004).

I thus start with examining what epistemic norms are and which role they might play in philosophical theories about science (Sect. 5.1). Then I distinguish different dimensions of normativity in philosophy of science. I clarify which relation exists between the normativity of a philosophical account and the fact that it includes claims about epistemic norms and in which sense descriptive-normative accounts are normative (Sect. 5.2). In Sect. 5.3 I argue that descriptive-normative projects 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 specifying in which sense my own analysis of reductive explanations in biology is normative and to which kind of philosophical projects it belongs (Sect. 5.4).

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)

Waters’ argumentation thus 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 and 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. It should for instance 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).[25] But 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).

5.2 Dimensions of Normativity

In Sect. 3.4 I have already distinguished two ways in which a philosophical account about a certain element of scientific practice X (e.g., reduction, concept of a gene, mechanistic explanation, etc.) can be normative: first, descriptive projects aim at understanding actual scientific practice and include only factual claims about X. Descriptive projects clarify what characterizes actual cases of reduction, how the concept of a gene is used in different biological fields, or what features mechanistic explanations in neuroscience have. But as I have shown in Sect. 3, even descriptive projects involve what I call *methodological normativity*. That is, they (often implicitly) rely on normative claims that concern the selection of those empirical information on the basis of which the philosophical account about X is developed. For instance, they presuppose normative assumptions about which examples should be conceived as paradigmatic, how the importance of an example for a certain field is estimated (if importance is coupled with success, how scientific success should be understood), and which statements of biologists should be conceived as correct or representative and thus should be included in the empirical basis of the philosophical account. Methodological normativity is inevitable, even though the normative claims that guide the selection of the empirical basis of an account are often left implicit.

Second, methodological normativity must be clearly distinguished from what I call *normativity in the strict sense*, which is characteristic of normative projects in philosophy of science. In normative projects normative claims about the subject matter X itself are made. For instance, the GRR model draws a picture of what theory reduction in the medical sciences *should* be; it does not represent how actual cases of theory reduction in this field in fact look like. Moreover, normative projects either contain no factual claims about X or these factual claims play no role in developing and justifying the account about X. But I will not take this to be a requirement for normativity in the strict sense as this would render impossible cases in which we have methodological normativity *and* normativity in the strict sense (as

---

[25] 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).

it might be the case in some descriptive-normative projects). The emerging picture can be summarized as follows:

Dimensions of Normativity in the Philosophy of Science
A philosophical account of a certain element of science X is normative if

(1) Methodological normativity (descriptive projects)

- it contains factual claims about X and
- it (implicitly or explicitly) relies on *normative claims about how to select the empirical information* about X on the basis of which the account is developed. Or

(2) Normativity in the strict sense

- it contains *normative claims about* X (and)
- [it contains factual claims about X, but they are not crucial for the development and justification of the account (only for normative projects).]

This picture can be further refined by considering the issue of social and epistemic norms in science. In which sense is a philosophical account that includes claims about social or epistemic norms normative and how does that relate to the two other dimensions of normativity? As a first step, let us consider claims about norms in science in more detail. Two kinds of claims can be distinguished: on the one hand a philosophical theory can be about social or epistemic norms, in other words, the norms themselves can be the subject matter X of a philosophical account. For instance, philosophers analyze the process of theory choice in science and discuss whether different epistemic norms (such as explanatory power, novel predictions, or simplicity) constitute plausible theory choice criteria.[26] Similarly, there is a huge debate about how social norms do and should influence scientific practice. On the other hand, the primary subject matter X of a philosophical account can be a certain element of science (apart from norms), such as reduction, mechanisms, or the gene concept, and the account includes claims about how this element X relates to certain epistemic norms. This is the kind of claims about epistemic norms Waters (2004) is concerned with. An example can be found in Craver's account of mechanistic explanation in neuroscience (2007a). His account does not only reveal the characteristics of mechanistic explanations in neuroscience; it also relates mechanistic explanation to an important epistemic norm, namely explanatory power, in so far as it includes claims about how to evaluate the adequacy of mechanistic explanations.

Another, even more important distinction concerns the *type of claim* that is made about either norms in science themselves or about the relation between an element of

---

[26] For an overview see, for example, Sankey (2013).

science and epistemic norms. These claims about norms can be either *factual* statements (e.g., “In scientific practice novel predictive success *is* accepted as a criterion for theory choice.” or “In molecular biology mechanistic explanations *are* assessed as adequate/good if they reveal which activities of which molecules bring about the phenomenon to be explained.”) or prescriptive, *normative* statements (e.g., “The ability of a theory to make novel predictions *should* be accepted as a theory choice criterion in scientific practice.” Or “Mechanistic explanations in molecular biology *ought* to be evaluated on the basis of whether they tell a complete causal story on the molecular level.”). In other words, either one can describe that a certain norm in fact is accepted in scientific practice (or how X relates to this in fact accepted norm) or one can adduce reasons for why a certain norm should be accepted in science (or for why X should be related to this norm in that way). The former amounts to a description, or better, to a critical reconstruction of existing norms, the latter is a justification of certain norms that need not be in fact accepted in science, but that should be accepted.

These difference give rise to a third respect in which philosophy of science can be normative:

(3) Norm-normativity

A philosophical account of a certain element of science X is normative if

(a) it contains factual claims about X (where X is a social or epistemic norm) or about how X relates to epistemic norms (*description of norms*) or
(b) it contains normative claims about X (where X is a social or epistemic norm) or about how X relates to epistemic norms (*justification of norms*).

Note that this third dimension of normativity is not completely distinct from the second dimension. Philosophical accounts that satisfy (3b) also satisfy (2). And factual claims about accepted norms in science (3a) also invoke methodological normativity, that is, they rely on normative claims about how to select relevant empirical information about norms in science. But still I think the third dimension of normativity is not dispensable as it highlights the importance of social and epistemic norms for normativity and promotes conceptual clarity.

Recall Waters’ view of conceptual analysis in philosophy of science discussed in Sect. 5.1. On the basis of the just introduced distinction between the description of norms (3a) and the justification of norms (3b) it becomes clear that Waters’ view is quite demanding. He requires a philosophical analysis for instance of the gene concept not only to contain a description of those epistemic norms that are actually accepted in science (which would be 3a). Instead, Waters argues that philosophers need to make independent, critical judgments, for instance, about the standards according to which the adequacy of gene-based explanations is to be evaluated (3b). In his words,

> 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.[27]

5.3 The Normative Part of Descriptive-Normative Projects

Let us come back to the third kind of philosophical projects from which this section started, jointly descriptive and normative projects. How can descriptive-normative projects be classified into these three dimensions of normativity? Which kind of normativity do they display? The answer to this question is not the same for all descriptive-normative projects and it depends primarily on how the normative part of the project is specified. One group of philosophers characterizes the normative part of their descriptive-normative project in a way that suggests that only a certain kind of methodological normativity (1) is involved here, or, if claims about epistemic norms are made, that these are only critical descriptions of accepted norms (norm-normativity, 3a). 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 methodological normativity only. The same is true for Wimsatt who argues 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). The point they are driving at is that philosophers should not merely passively image scientific practice, but rather actively and critically reconstruct the practices and insights they find in contemporary science. As I have argued in Sect. 3, such a critical reconstruction involves making normative claims about what should be regarded as important case studies and as successful science. In line with this, 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).

Other 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

---

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

scientific practice. Rather, they stress that their project comprises a *justification* of certain epistemic norms, which is supposed to be *independent* from their critical description.[28] 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, norm-normativity (3b). 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 (second dimension of normativity, normativity in the strict sense) 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).

5.4 A Dilemma

Craver's claim that the normative part of his project is modeled upon or informed by the descriptive part poses the question of what relation exists between factual and normative claims in descriptive-normative accounts. 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.[29] 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

---

[28] 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.

[29] A similar idea can be found in Schindler (2013).

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 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.[30] 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

---

[30] Schindler also recognizes this problem and proposes an *a priori* justification of certain norms. We should ask whether it is "*plausible* that, counterfactually, the ends [of science] *would* be achieved if certain norms… *were* to be respected" (2013, 4150).

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 from normative ones.

In conclusion, jointly descriptive and 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 Sect. 2.

5.5 How Normative My Account Is

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 displays *methodological normativity* (first dimension of normativity, Sect. 3.4). My account deviates from a mere description of different examples of reductive explanations in biology and of how biologists understand the term ‘reductionism’ 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 (Chap. 5, Sect. 3.1). 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 (recall Sect. 3.2).

But 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 (Sect. 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 will address in Chap. 6. 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; see also Kaiser 2011). 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 (such as explanatory success), a third criterion of adequacy must be added[31]:

Criterion of Norm-normativity

If the aim is to understand biological practice, an account of reduction in biology 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.

---

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

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. It thus exhibits *norm-normativity* (third dimension of normativity). But since my account does not include the independent justification of epistemic norms such as explanatory success it displays norm-normativity only as a critical description or reconstruction of such norms that are in fact accepted in contemporary biological practice. This clearly distinguishes my account from normative projects in philosophy of science as well as from descriptive-normative projects that seek also an independent justification of norms.

6 The Relevance of Philosophy to Science

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). The utility of philosophy to science requires, on the one hand, that philosophy transcends being a mere description or image of 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 their advice should be "contextual and sensitive to feedback, not a priori pronouncements offered ex cathedra" (Wimsatt 2007, 27). Philosophy of science thus 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 this objection 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). On this basis the following fourth criterion of adequacy for a philosophical account of reduction can be formulated:

Criterion of Potential Usefulness
If the aim is to understand biological practice, an account of reduction in biology 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 normative projects in philosophy of science (recall Sect. 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 descriptive-critical (or descriptive-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 actual 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 normative account that involves no empirical information about actual science is just too far away from actual scientific practice to be potentially relevant to science. But the vast majority of accounts that are developed in philosophy of science and in philosophy of biology, in particular, are not of the normative kind. With respect to these accounts I think that potential utility is an important criterion of adequacy. I thus agree with Wimsatt that philosophers of science should not only be "theorists of reason", but also "*therapists of reason*" (2007, 27).

7 Interim Conclusion

In this chapter I disclose the meta-philosophical assumptions that lurk behind my analysis of reductive explanation in biology. The aim of my analysis is to provide an understanding of reduction as it occurs in biological practice. This aim indicates that I am pursuing what I call a descriptive-critical project, contrary to a normative project that aims at constructing an ideal of what reduction should be. Given the aim of understanding reduction in biological practice, I argued that any adequate account of reduction in biology must satisfy four criteria:

Criteria of Adequacy
An account of reduction in biology is adequate iff

(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 (**Descriptive Adequacy**),

(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 (**Balancing Specificity Against Generality**),

(3) it clarifies in which sense reductions contribute to or hinder the realization of relevant *epistemic norms*, for instance, explanatory success (**Norm-normativity**), and

(4) it is *potentially useful* to contemporary biological practice and to the broader society (**Potential Usefulness**).

These criteria specify the empirical basis, on which I will develop my account of reduction: actual cases of reduction that are paradigmatic and important (1*a). They elucidate the philosophical methodology of descriptive-critical projects like mine: accounting for the empirical basis that often is diverse (1*a and 2), explicating implicit assumptions (1*b), and establishing coherence (2). These criteria also formulate certain virtues that the resulting account of reduction must have: it should be as universal as possible (2), it should clarify how reduction is related to major epistemic norms (3), and it should be potentially useful (4).

Another major result of this chapter concerns normativity. Even if my project does not belong to the sort of projects that I have characterized as normative projects this does not mean that my analysis of reductive explanation is completely free of normativity. The kinds of normativity involved in my analysis are just different from the kind involved in normative projects. As a descriptive-critical project, my analysis relies on claims that are methodologically normative (normative claims about which empirical information is relevant) and because of criterion (3) my analysis also displays norm-normativity (it entails claims about epistemic norms). But since my account does not make normative claims about reduction itself it is not normative-in-the-strict-sense (as normative projects are).

Even though this chapter primarily serves to illuminate the goals and methodology of my analysis of reduction, the views I develop and, in particular, the four criteria of adequacy that I propose point beyond my specific project of analyzing reductive explanation in biology. These criteria can easily be adapted such that they apply more broadly to philosophy of science in practice.

Chapter 3
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 and Sahotra Sarkar 2006, 697)*

Contents

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 sixteenth and seventeenth 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 1970s. 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 50 years so far. But this

M.I. Kaiser, *Reductive Explanation in the Biological Sciences*, History, Philosophy and Theory of the Life Sciences 16,
DOI 10.1007/978-3-319-25310-7_3

introduction is not a mere overview. Rather, I present what I take to be the most crucial lessons one should learn from the previous 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 (Sect. 1). Second, if you seek to understand what reduction in actual biological practice is, you should focus on epistemic, rather than on ontological issues. But you should keep in mind that questions about ontological reduction and about the interrelations between epistemic and ontological issues, nevertheless, are relevant (Sect. 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 (Sect. 3). And since, fourth, it is time to move beyond Nagel's inadequate model of theory reduction (Sect. 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; 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 follow this trend. 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]

The project of understanding reduction is not only different from the project of disputing reductionism; I think the former is also *prior* to the latter. The reason is that any *convincing* argumentation in favor of a particular version of 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 and 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 what it means to reduce an entity to another.[3] As I will explicate 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 biological object to be explained (explanandum) is reduced to the description of the explanatory relevant factors (explanans) (see also Chap. 4, Sect. 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

---

[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.).

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

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 Chap. 6.

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.

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:

Criteria for Distinguishing Versions of Reductionism

(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 reductionist position with a relatively broad scope (e.g., "All biological phenomena can be explained in physicochemical terms.") do not confine their claims to reductions that are indeed feasible in contemporary science (since this would make their claims false). Rather, they argue that the impracticability

---

[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).

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 (Dupré 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).

A second criterion for telling apart different versions of reductionism is the difference between eliminative and retentive reductionism:

Criteria for Distinguishing Versions of Reductionism

(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 often 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 and 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.

Third, reductionist positions can also be distinguished by their scope:

Criteria for Distinguishing Versions of Reductionism

(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)?

---

[5] As I will argue in Chap. 4, 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 Sect. 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 and Sarkar 2006, 698–700 and Wimsatt 2006a, 457.

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 and Wagner 2001; Frost-Arnold 2004).

Finally, different versions of reductionism include different claims about how far reduction goes, that is, what the reducing locus is:

Criteria for Distinguishing Versions of Reductionism

(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 Chap. 6, Sect. 2.4). 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 Chap. 4, Sect. 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. 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 Chap. 6, Sect. 2.4.3). 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.

But still, my proposal of how to understand reductive explanation in biology leaves many options 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]

To conclude, 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 (see Sect. 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

---

[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 Chap. 4, Sect. 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).

of applying reductive strategies in certain research fields – are of peculiar interest. Most philosophers of biology quite rashly agree on ontological reductionism (in shape of a token physicalism, see the next section) 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). Or 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 Sect. 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 (Sect. 2.1) and, then, I try to shed light on the connections that exist between these two issues (Sect. 2.2). I conclude by showing why I nevertheless think that epistemic issues are more relevant to biology than ontological ones (Sect. 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

---

[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".

(e.g., between theories from 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 and 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 (Sect. 2.1.1) and by epistemic reduction (Sect. 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 Sect. 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 (even if not all) to be learned from philosophy of mind.

But 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 Sect. 2.2.

[10] This is not to say that there exist no close connections between ontological and epistemic issues concerning reduction (see Sect. 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 and 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. But 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 this is the way the term 'ontological reduction' *must* be understood. This 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 a relation of *constitution* or localization. In what follows I will amplify these claims. But as an extensive debate and an immense bulk of literature in philosophy of mind stand behind each of these claims,

---

[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 2004b, 148) as expressing an *ontological* and not an epistemic claim.

I will focus on 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; 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 2004b, 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 and 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 and 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 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 (for an extensive discussion of Nagel's model see Sects. 3.1 and 4). 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–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 Sects. 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

---

[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).

physicalism is advocated (e.g., Pereboom and Kornblith 1990; Pereboom 2002; Baker 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 seems 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 and 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 and Bennett 2011). Whatever these specifications amount to, what interests me at this point, is that there seems to exist an important difference between the two options outlined above,

---

[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 in philosophy of biology 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.

namely token physicalism and supervenience physicalism.[22] As I have already indicated, 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.

---

[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 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).

The answer that has dominated the debate in philosophy of mind to this day 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 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 viable alternative seems to be to appeal to the relation of *constitution* to specify the relation of reduction. 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 and 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 and 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. But one need not interpret this phrase in this way. It might also be interpreted as expressing the fact that for

---

[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.

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 denies Schaffner's claim that identities are required for successful reductions (1967). Wimsatt argues that, instead, "localizations" (2006a, 456; see also Bechtel and Richardson 2010, xxxii) suffice, too, because they "preserve all relevant spatiotemporal properties of identities" (Wimsatt and Sarkar 2006, 700). He states that localization is distinct from identity, more precisely that localizations are "logically weaker than identities" (2006a, 456). Wimsatt's few remarks, however, 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 Chap. 6 I will argue that part-whole relations 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 biological 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, propositions, 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 Sect. 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 Sect. 4). Moreover, at the beginning the debate focused on quite *global* cases of epistemic reduction, such as 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). Many 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 and Casadevall 2011, 1401). Furthermore, this was also the way the notion of epistemic reduction(ism) was introduced in philosophy of biology (Ayala 1974, ix).

But to think about epistemic reduction by focusing on theories, on formal issues, and on global cases of reduction is not the only route that has been taken so far and, as I will argue in Sects. 3 and 4, it is not the most promising way to understand what epistemic reduction in biology 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 Sect. 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

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 and Love 2008, Section 1; Schaffner 1993, 515), yet to recognize also the relations that hold between ontological and epistemic reduction. But how exactly are ontological and epistemic reduction related to each other?

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 (Rosenberg 2006), or even more radically, that ontological and explanatory reduction fall together (Brandon 1996). Although Brandon acknowledges the difference between methodological, ontological, and explanatory reductionism, he argues that it "would be tedious to continually make this threefold distinction" and that the "very tight connection between the explanatory and ontological versions [of reductionism]" (1996, 180) allows him to lump together ontological and explanatory reduction(ism) under the single label 'ontological reduction(ism)'. Rosenberg also assumes a tight connection between ontological and explanatory issues. 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 more fundamental entities and processes (such as molecules or physical particles) – 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 turns out to be unconvincing at several points. In short, according to Rosenberg biological explanations are adequate only iff

(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 and 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 the principle of natural selection (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 provide a maximally complete description of the causal process that underlies the explanandum

---

[26] The details of Rosenberg's view about reduction and explanation will be fully elucidated and critically examined in Chaps. 4 and 5.

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, in his view 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, such as the question being asked or the research interests of biologists. Rosenberg claims that a pragmatic, erotetic view of explanation results in subjectivism, which is why he rejects it (2006, 35f, 44, 179f).

To sum up, Rosenberg is right that 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 Chaps. 4 and 5 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 objections 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 epistemic-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 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 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. Briefly speaking, ontological reduction does *not* enforce epistemic reduction.

2.2.2 Alternative Relations

But caution is needed here. Even if it is wrong that you cannot have ontological reduction without at the same time having explanatory reduction, this is not to say that there exist *no* connections between ontological and epistemic reduction at all. There are important connections, and 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 Sect. 4). But 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] But the reductive explanation of a particular biological phenomenon still requires token reduction (understood either as relation of identity or of constitution/localization). For instance, if an individual ATP synthase would not be

---

[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.

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 biological object or system *Y* that refers to certain biological parts of *Y* and to certain interactions between these parts would be inadequate if *Y* were not in fact composed of these biological 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. The term 'ontic' does not express that the object of my analysis is ontological reduction. Rather, my analysis is about a specific kind of epistemic reduction, namely explanatory reduction (i.e., reductive explanation). My account of explanatory reduction is ontic (analogous to the ontic account of explanation; Salmon 1989, 1993, 1998) since it includes the central claim that what determines the reductive character of an explanation is whether it represents certain relations that *exist in the world* (e.g., '*X* is located on a lower level than *Y*' or '*X* is internal to *Y*', i.e., '*X* is a biological part of *Y*'; for further details see Chap. 6, Sect. 6). These relations in the world that determine the reductivity of explanations are not the same as the relations of ontological reduction, but related to them.[31]

Methodological reduction implies various ontological issues, too. In Chap. 6 I identify two respects in which the application of reductive methods (as decomposition and separating a biological object from its environment) can fail and does not yield reductive explanations. First, a biological object or system *Y* can exhibit complex forms of interactions with its environment that make it difficult or even impossible to separate *Y* from its environment. Second, several biological objects or systems are far from being "aggregative systems" (Wimsatt 1986), but rather are highly organized and "functionally integrated" (Bechtel and 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

---

[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'.

(see Chap. 6, Sect. 4.2.2). Since non-aggregativity, non-decomposability, and non-separability are all features of biological objects or systems *in the world*, the failure of reductive methods can be traced back to ontological issues.[32]

2.3 Why Epistemic Issues Matter Most

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 good 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). But 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 interacting chloroplast membranes, thylakoids, ribosomes, etc.' (for further evidence see Chap. 6).[34]

---

[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.

[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.

The questions biologists are rather occupied with are methodological and explanatory issues. They worry about 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 an object'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 of 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 and 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 (recall Chap. 2).[35] My project of analyzing reductive explanations in biology lines up with this tradition since it is also focused on cases of epistemic reduction. But the fact that epistemic questions are the issues that matter most does not imply that questions about ontological reduction are irrelevant. Accordingly, in this section I hope to have shed some light on what the kind of ontological reduction(ism) is that underlies the debate about epistemic reduction(ism) in the philosophy of biology (Sect. 2.1), and which connections exist between, on the one hand, epistemic reduction and, on the other hand, ontological reduction or ontological issues in general (Sect. 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.

---

[35] By the way, just as philosophy of biology can learn much about ontological reduction from philosophy of mind (see Sect. 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).

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 (such as an organism showing a certain behavior) be said to be identical to or constituted by certain physical objects and processes?" 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 entities. In this case the question could be specified for instance as follows: "Can the knowledge about or the representation 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) manner. Rather, it is either formulated as a question about the relation between *theories* ("Can a biological theory about a certain phenomenon be reduced to certain 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. He distinguished 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 many philosophers. Some authors add new subcategories (e.g., theory and explanatory reductionism as subtypes of epistemic reductionism; Sarkar 1992; Brigandt and 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; cf. Sect. 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 Sect. 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 not only by drawing on results from the previous debate, but also by breaking new ground. For example, I spell out the notion of methodological reduction by focusing on reductive methods rather than on the normative claim of a methodological reductionist (Sect. 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 (Sect. 3.3; see also Chaps. 4 and 5). But first, let us start with examining the most traditional 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 does not 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 the 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 and 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 also 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 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

---

[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.

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

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, b, 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 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).

However, despite these modifications, at that time hardly anybody questioned that Nagel's model is, in general, the adequate way of thinking about epistemic

---

[39] Notable exceptions are Ruse (1976) and Schaffner (1967, 1969).

[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).

reduction. That is, most philosophers accepted the following two *core assumptions* of Nagel's account:

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 and 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

---

[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.

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 and Dietrich 2006; Darden 2005; Darden and Maull 1977). Contrary to them, I think it is important to 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 new account of what epistemic reduction in biology is, not the removal of the concept of reduction from philosophy of biology altogether.

In Sect. 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 Chaps. 4, 5, 6 and 7, 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 (Sect. 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 Sects. 3.2.2 and 3.2.3 I provide first steps in developing such a concept. In Chap. 6 I pick up on this task again, identify and analyze 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

---

[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").

and molecules" (Ayala 1974, viii). Thus, proponents of methodological reductionism endorse a *normative* claim or recommendation about how research in the biological sciences should be done. Their argument is that biological research should be pursued in a reductionist fashion because this is the most successful way to pursue it.

Typically, the phrase 'pursuing 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… 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 understanding of methodological reductionism. He accuses methodological reductionists of practicing "*wannabe reductionism*" (2006a, 445; my emphasis). The reproach is that methodological reductionists claim that one should pursue reductionism, but they 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, fn. 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 and Sarkar 2006, 697). Since we want philosophy to be connected to what scientists actually are doing and to be potentially useful for practicing scientists (recall Chap. 2, Sect. 6) this gap needs to be closed. In the subsequent sections I thus first spell out the notion of a reductive method (Sect. 3.2.2) and then specify the relation between reductive methods and reductive explanations (Sect. 3.2.3). Finally, I address the putative objection that reductive methods, not reductive explanations, are the appropriate target of an analysis of epistemic reduction (Sect. 3.2.4).

---

[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).

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. 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.

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 and 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 in understanding what methods are there are two extremes: on the one hand, one could regard reductive methods as specific techniques or, on the other hand, one could identify methods with global stances. Consider first the option to characterize reductive methods as small-range methodological elements, such as laboratory *techniques* (e.g., PCR, Southern blot, screening, etc.). The disadvantage of this option 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 option, that reductive methods are global elements, such as entire reductionist *approaches*, *stances or attitudes*, is at first sight attractive because this is what some biologists mean by 'reductionism'. For example, in cancer research the putatively misguided "reductionist stance" (Soto et al. 2009, 3) is opposed to the integrative, "organicist approach" (Soto et al. 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. But 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

---

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

in succession), is what several biologists seem to regard as good research. They do not argue against the application of reductive methods in their research field *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 and Appenzeller 1999, 79). In their words, "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).

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 2004a, 1016), which is also known as "*decomposition*" (Bechtel and Richardson 2010, 23; my emphasis), "(functional) analysis" (Ayala 1974, vii; Cummins 1975, 1983; 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 biological object, namely the cell, into its biological parts (e.g., ribosomes, the genome, amino acids, m- and t-RNAs, splicing enzymes, etc.) and to study the properties or behaviors 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 object or system as well as different biological objects or systems can be decomposed differently, that is, into different kinds of biological parts (cf. Kaiser forthcoming a) and into biological parts that are located on different levels of organization (cf. Kaiser manuscript a). 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 biological parts and interactions/activities are taken to be unproblematic and relevant to the explanatory interests and purposes of a given scientific field. Even though the decomposition of

---

[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).

many biological objects identifies molecules and their interactions as biological parts, the decomposition need not bottom out at this fundamental level (Chap. 6, Sect. 2.4.3). 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 73–93 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 objects or systems exhibiting a particular behavior (e.g., the transformation of light energy into chemical energy in chloroplasts) can be decomposed into two kinds of entities. 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 processes, operations, interactions, activities, or occurrents in general (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 relations and on objects for a long time. However, my account of biological parthood (Kaiser forthcoming a) reveals that biological processes, in which biological objects are involved, and the temporal and relevance relations between these processes are as important for part-whole relations as objects and the spatial relations between them (see also Chap. 6, Sect. 1.1).

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 biological object or 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 "mapping the operations into which the overall function of the mechanism

---

[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.

[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 Chap. 6 I come back to several reductive methods again and discuss them in more detail. 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 point out 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 in a too negative picture of reductive methods – as David Hull and Marc van Regenmortel put it,"[r]eductionistic science is not all bad" (2002, 12). Moreover, 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 object or system by decomposing it into biological 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 be reductive?

---

[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 Chap. 6. 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.

I agree that reductive methods and reductive explanations are two elements of biological practice that are closely tied to each other. This is the reason why my account of explanatory reduction also mentions reductive methods. But yet, 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, reductive explanations and reductive methods are closely connected, but not the same. More importantly, there are clear cases where the two can come apart: cases, in which the application of a reductive method does *not* result in 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 object or 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 lead to false results, that is, it can lead to the development of reductive explanations that are *inadequate* – 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 (more on this in Chap. 6). 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; Brigandt and 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*. To conclude, the reductivity of explanations is *partially independent* from the reductive character of the applied methods.

In many cases, however, 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*

---

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

of methods: reductive methods together with non-reductive ones (e.g., Mitchell 2009; Brigandt and Love 2008, Section 5; Bechtel and 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 and Appenzeller 1999, 79; see also Sect. 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]

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, b, 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. In his view, 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 he adduces 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

---

[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.

useful, he thinks that philosophers need to shift their attention to the way the "investigative strategies" (2008, 257) in genetics have changed.[58]

I endorse Wimsatt's and Waters' careful consideration of biological practice. 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' thesis about the distinguished status 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 as genetics has developed or been transformed over time. My analysis is concerned with interlevel reduction, not with successional reduction (cf. Sect. 3.4).

Second, it seems to me that from the perspective of the biological sciences reductive explanations play an equally significant role in biological research practice as reductive methods. Reductive explanations thus 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, my focus on reductive explanations is not only due to my personal interests. I also think there exist good reasons for a philosophers of biology to turn their 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. Reductive explanations thus constitute a more suitable target of philosophical analysis. However, this does not imply that *no* constrained and transdisciplinary methods can be found in scientific practice, which are adequate objects of philosophical investigation. Nor does it imply that we don't need analyses of methodological reduction in order to get a comprehensive view of epistemic reduction in biology.[60]

---

[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.

[60] 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 and 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. Hence, even if Nagel's model is primarily a model of theory reduction it could be read also as a model of explanatory reduction.[61]

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 and 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 Chap. 4, Sect. 2; my analysis of explanatory reduction is presented in Chap. 6).

3.4 Successional vs. Interlevel Reduction

This section (Sect. 3.4) is not fully in line with the other three (Sects. 3.1, 3.2, and 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

---

[61] 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 Sect. 3.1) and reconstruct reduction primarily as a relation between explanations or between parts of an explanation.

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.[62]

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.[63] 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).[64] 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 (recall Sect. 1). Interlevel reduction is a relation between two theories from different domains that typically are accepted at the same time. The reducing and the reducing theory typically have a different area of application: the reducing theory makes assertions about objects that are located on a lower level of organization than the objects the reduced theory makes claims about. That is, interlevel reduction relates a higher-level theory to a lower-level theory, as for instance the putative reduction of Mendelian genetics to molecular biology.[65]

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).[66] One should notice that the fact that there exist no (or only very few) cases of interlevel theory reduction in biology does not show that the conceptual distinction between successional and interlevel reduction is misleading in the context of theory reduction.

---

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

[63] 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.

[64] 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.

[65] 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.

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

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.[67] 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. Questions about successional explanatory reduction would then be related to the development of explanations in one domain across successive historical periods of scientific change.[68]

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. This decision may provoke 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 of 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 commit me to analyze successional reduction rather than interlevel 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 Chap. 6) 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 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

---

[67] These two kinds of interlevel reduction correspond to the two types of explanatory reduction characterized in Chap. 4.

[68] This might be similar to what Rosenberg (2006) has in mind (see Chap. 4, Sect. 1).

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 I first identified the main lessons one should learn from the previous debate about reduction(ism) in biology. These lessons at the same time delineate and justify the path I am going to take in answering the above question: in Sect. 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 working biologists. In Sect. 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 Nagel's influential model of theory reduction and to turn our attention to the search for an adequate model of explanatory reduction in biology (cf. Kaiser 2012).

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 Sect. 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).[69] Nagel's particular view of the *relata of reduction* encounters a serious objec-

---

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

tion, namely that the relata of reduction he envisions do not exist in biological practice. Neither do there exist 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 of 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.[70] 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, 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 (recall Sect. 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.[71] 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).

---

[70] 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).

[71] 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 Sect. 3.1:

Core Assumptions of Nagel's Account

(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 (Sect. 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 (see also Kaiser 2012). 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 (Sect. 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.[72] 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 et al. (2010) want to convince us not to be afraid of Nagelian

---

[72] 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 and 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.

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).

Likewise, Bickle (1998) adheres 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 the present 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 that the relation of reduction is a relation of *logical derivation* (recall the previous section). But 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 (see also Kaiser 2012). To put it in a nutshell, first, biological practice shows that, in general, theories are not the only (and perhaps not the most) important element of scientific practice. Second, biological practice 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 formal features of theories the syntactic view fails to capture what biological theories in fact are (i.e., 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, 2013). 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 that other epistemic units (such as models, descriptions of mechanisms, fragments of theories, etc.) play in explanation, prediction, discovery, and manipulation in the biological sciences.

The latter criticism can be expanded to a more general objection that applies to any philosophical 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[73] 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 and 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.

---

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

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 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 and Casadevall 2011, 1401)[74]

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. Nowadays, the majority of philosophers of biology hold this view.[75] 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 Rosenberg, who defended Nagel's approach for a long time (1985, 110; 1994, 22), demands that "the question of what reductionism was in the post-positivist past" must 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 that 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

---

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

[75] 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).

"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; my emphasis) of what a *complete reduction* would look like. He admits (as he has done 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 of reduction (or, at least, as a "secret hope or end" Wimsatt 1976a, 685). Hence, Schaffner remains one of the very few philosophers of biology who resists 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 typically is.[76]

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 and 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 Chap. 6, Sect. 2) and since some of them encompass black boxes that are not filled in yet (Machamer et al. 2000, 18; Schaffner 2006, 397). Also the reductive methods that are applied in the biological sciences are not global stances (recall Sect. 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.

---

[76] 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 Chap. 6 I analyze several of these examples, each of which could be adduced as empirical evidence for my theses here.

Second, the Nagelian model of theory reduction has been closely connected with the so called "*layer-cake*" *picture of science* (Oppenheim and Putnam 1958).[77] 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 theory from a lower-level one.[78] 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 actual biological practice. 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 Chap. 5, Sect. 4.4). Several biological phenomena cannot be reductively explained and even reductive explanations are not restricted to fundamental level explanations (see Chap. 6, Sect. 2.4.3).

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.[79] 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 (which is why my own account of epistemic reduction will highlight these issues and will be ontic in character).

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 (Sects. 4.1 and 4.2), Nagel's general way of thinking about epistemic reduction in terms of

---

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

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

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

theories and their logical relations still remains inadequate with regard 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 (Sect. 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 actual typical cases of epistemic reduction in biology 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 concerning epistemic reduction (Sects. 4.3.2 and 4.3.3). Hence, the fourth lesson we should learn from the previous debate is that it is time to move beyond Nagelian reduction and to shift our attention from theory reduction to reductive explanations and from formal to substantive issues.[80]

5 Interim Conclusion

The goal of this chapter is 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.

The four lessons one should draw from the previous debate are the following: *first*, before you discuss whether reductionism or antireductionism is true you should seek to understand what reduction is (Sect. 1). 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 (anti-)reductionism 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.

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, not ontological issues, are the ones that matter most (Sect. 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 (Sect. 2.2).

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

---

[80] 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)

about ontology than “We’re all ontological reductionists. Case closed”. This is why, in Sect. 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 and of epistemic reduction I then turned to the question which relations exist between the two issues (Sect. 2.2). The first answer I gave was negative. 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. As my own analysis of explanatory reduction will display, the availability of an adequate reductive explanation of the behavior of a particular biological object or system presupposes that certain relations and facts exist in the world (e.g., level relations, part-whole relations, 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 (Sect. 2.3). The biological literature (which will be analyzed in more detail in Chap. 6) 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 (Sect. 3.1), methodological reduction (Sect. 3.2), and explanatory reduction (Sect. 3.3). Furthermore, it is important to tell apart cases of successional reduction from cases of interlevel reduction (Sect. 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. I pursued two goals in this section:

My first aim is 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.

I argued that common formulations of methodological reductionism do not sufficiently specify. I undertook first steps to fill this gap 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 Chap. 6). Moreover, I revealed that the usage of reductive methods can but need not result in reductive explanations. 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.

My second goal is to adduce reasons for why it is most promising to analyze epistemic reduction in biology by focusing on explanatory reduction, and not on methodological reduction or on theory reduction. I stated that methodological issues concerning reduction are significant and do play some role in my analysis. But 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 is to show that, even if one admits that Nagel's original model can be improved in several respects (Sects. 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 and which are particularly insignificant in the context of epistemic reduction. Even a refined version of Nagel's model fails to capture real cases of epistemic reduction in biological practice 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.

Chapter 4
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)*

Contents

In the preceding chapters I revealed the methodological preliminaries of my analysis of reduction in biology (Chap. 2), I introduced the previous reductionism debate and, thereby, pointed to the direction in which my analysis will proceed (Chap. 3). The methodological procedure by which I develop my account of reduction and the kind of account I seek should now be 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 biological research as well as their strengths and limitations (i.e., the conditions under

M.I. Kaiser, *Reductive Explanation in the Biological Sciences*, History, Philosophy and Theory of the Life Sciences 16,
DOI 10.1007/978-3-319-25310-7_4

which they are adequate). But before this analysis can start (Chap. 6), it is necessary to 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, the notion of explanation that underlies my analysis must be carved out (Chap. 5).

The task of this chapter 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. This chapter thus 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. 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; recall Chap. 3, Sect. 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 particular view of explanation (e.g., Sarkar 1998; see also Chap. 5). 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 regards 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 notion of explanatory reduction he accepts. 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 (Sect. 1). After reconstructing the major aspects of his defense of Darwinian reductionism (Sect. 1.1), I reconstruct the notion of explanatory reduction he presupposes in his argumentation (Sect. 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 (Sect. 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 (Sect. 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 (Sect. 2.1). Afterwards, I consider Sarkar's work on reductive explanations in genetics and molecular biology (Sect. 2.2). I explicate what it means that Sarkar calls his account substantive (Sect. 2.2.1) and why he emphasizes its neutrality with respect to any account of explanation (Sect. 2.2.2). Then I focus on the core of his account, namely his three criteria of reductivity (Sect. 2.2.3). I conclude by revealing the major shortcomings of his analysis (Sect. 2.2.4). In Sect. 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 (Sect. 2.3.1) I scrutinize their three aspects, namely intrinsicality, fundamentality, and temporality (Sect. 2.3.2), and clarify in which sense these aspects constrain the reductive character of a biological part-whole explanation (Sect. 2.3.3). Finally, I disclose the difficulties their analysis encounters (Sect. 2.3.4).

In the conclusion (Sect. 3) I summarize what I take to be the most important insights and shortcomings of these previous two 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 (recall Chap. 3, Sect. 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] One of my tasks in this section will thus 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 (Sect. 1.1.1). After that I characterize the version of explanatory reductionism Rosenberg advocates (Sect. 1.1.2). To do this, I use the classification I introduced in Chap. 3, Sect. 1. Then I give a rough outline of the general line of argumentation by which Rosenberg defends reductionism (Sect. 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 (Sect. 1.2). Finally, I point out several difficulties Rosenberg's notion of explanatory reduction faces (Sect. 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] 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).

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

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.

Rosenberg has, however, not always been a defender of reductionism in biology. Rather, this is a novelty of his 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 of 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 (more details in Sect. 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 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. Hence, 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 (recall Chap. 3, Sects. 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 9 years before: above the level of molecules nature is very complex. This

---

[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.

[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).

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 Dupré 1993) in his 1994 book he adopts an instrumentalistic, anti-realistic interpretation of biological theorizing. 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 from the failure of Nagelian reduction: "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 abandonment of Nagel's model 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:

> [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, reductionism as a claim about explanations, and reductionism 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. But Rosenberg’s overall argumentation suggests that he better should have dropped the term ‘more’ (and he does so in other 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)

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 Sects. 1.1.3 and 1.2).

---

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

Let us now take up the classification I introduced in Sect. 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 biological 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. But 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 biological field. Fourth, Darwinian reductionism seems to presuppose a very *restricted locus* of explanatory reduction, namely the level of molecules. Rosenberg claims 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 Sects. 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 (Rosenberg subsumes any biological field except molecular biology under the label 'functional biology'). Moreover, 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 anti-reductionists.

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, why the trait "spots on wings that resemble owl eyes" was established in populations of *Junonia coenia* over time, that is, 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 Chap. 5). 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*. 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 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. But, as I have mentioned in Sect. 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,

---

[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 Chap. 5, Sect. 4.4.

[7] More about Rosenberg's view on explanation can be found in Chap. 5.

[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 Sect. 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.

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 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 4.1 illustrates the difference between these two readings.

In sum, Rosenberg treats explanatory reduction as a relation between an ultimate and a molecular proximate explanation of the same phenomenon. The process of

---

[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.

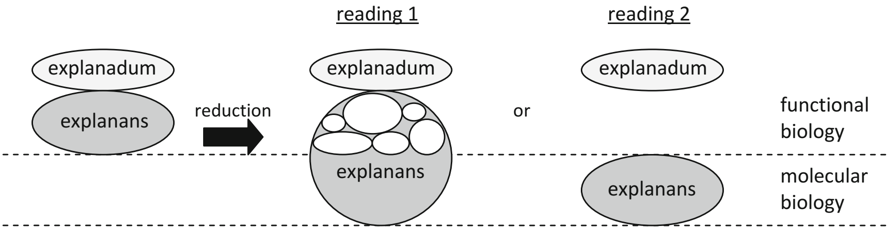

Fig. 4.1 Rosenberg's notion of explanatory reduction

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.

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). In this section, however, 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. There are three additional shortcomings of Rosenberg's perspective that I want to draw attention to here.

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). But 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 (recall Chap. 2, Sect. 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 Sects. 3.1 and 4 of Chap. 3. What carries most weight is that Rosenberg does not pay much attention to what epistemic reduction in practice is, but rather constructs 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 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, fn. 6)

---

[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).

[12] This claim is supported by the various examples of paradigmatic and important cases of explanatory reduction I discuss in Chap. 6.

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

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 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.

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 Chap. 5, Sect. 4). 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.[14]

In conclusion, 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. Rosenberg's account thus fails to capture what epistemic reduction and what explanation in contemporary biological practice really are.[15] 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.

---

[14] 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 Chap. 6.

[15] The empirical evidence for this claim can be found in Chap. 6.

2 Second Perspective: Individual Reductive Explanations

To regard explanatory reduction as a relation between two explanations is not the only available option. An alternative perspective is to consider *individual* biological explanations and the constraints of their reductive character. Those who adopt this perspective 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 Chap. 6, 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 representation of the phenomenon or behavior of a biological object or system to be explained) and the explanans (i.e., the representation of the explanatorily relevant factors). 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 and to mechanistic explanations.[16] That is, part-whole and mechanistic explanations are treated as paradigmatic cases of reductive explanations (more on this in Chap. 6, Sect. 5).

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 I can take up in my own analysis of explanatory reduction. The central question to which I hope to find some viable answers is: 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 (Sect. 2.1), namely Kauffman's paper on part-whole explanations (1970) and Wimsatt's claims about reductive explanations (1976a, 2007).[17] 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 (Sect. 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 (Sect. 2.3). I conclude by summarizing the fruitful insights into the features of reductive explanations that have been offered so far (Sect. 2.4).

---

[16] 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).

[17] 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 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.

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"). But 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: 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 developed.[18] 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 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. Part-whole explanations thus 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 so-called New Mechanists (e.g., Machamer et al. 2000; Glennan 2002; Craver 2007a; Bechtel 2006, 2008) that mechanistic explanations

---

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

refer to both, to entities (i.e., the objects that are spatial components of a mechanism) and to what the entities are doing, that is, which interactions, activities, or operations they engage in. My own account of biological parthood (cf. Kaiser forthcoming a) is also similar to Kauffmann's claim. I argue that part-whole relations in the biological realm are not only determined by spatial relations between objects but also by temporal and relevance relations between processes.

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, Kauffmann argues that new information about the parts of a system may 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 emphasized by the New Mechanists (e.g., Craver and Darden 2001, 119–123; Bechtel 2006, 28–33).[19]

All in all, Kauffman's paper presents an important analysis of the nature of part-whole explanations and of the process of developing part-whole explanations via decomposition (recall Chap. 3, Sect. 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, what it is that makes part-whole explanations *reductive*. Kauffman's ideas thus do not bring forward 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 impressive 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 won't review 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 Sect. 4 of the previous chapter. What interests me at this point is whether Wimsatt's paper yields any insights into the features of reductive explanations in the biological sciences.

---

[19] 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.

Due to the fact that Wimsatt's primary focus is to criticize Nagelian accounts of reduction, only a few comments on what makes biological 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: with explanations that show how the phenomenon is a product of causal interactions at its own level and with explanations that show how the phenomenon is a product of causal interactions 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 (Sect. 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 (Sect. 2.2.1) and, on the other hand, his claim that an account of explanatory reduction must be independent from an account of explanation (Sect. 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 entire book (1998).

The distinction between formal and non-formal accounts of epistemic reduction, however, is not new. Wimsatt, for instance, characterizes his approach as "functional" (1976a, 675, 700) and contrasts it 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).[20]

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*.[21] 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 (recall Chap. 3, Sect. 2). I think it is important to emphasize that this is *not* what it means. Even if Sarkar focuses on

---

[20] 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.

[21] 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).

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 Chap. 6, Sect. 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 Sect. 2.2.3). These criteria (i.e., fundamentality, abstract hierarchy, and spatial hierarchy) can be formulated as relational predicates: 'is fundamental to', 'is located on a lower level of an abstract hierarchy than', and 'is located on a lower spatial level than'.[22] 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).

To see this 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 biological object or system, whose behavior is to be explained, is represented as being located on a *lower spatial level* than the objects and processes referred to in the explanans. 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 thus assumed to determine the reductive character of the explanation. This is what I think makes Sarkar's account substantive (or ontic).

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).

---

[22] 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).

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 (such as 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.

(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 in the sense that 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). But it remains obscure what this wider notion of derivation is and whether and how it can be distinguished from the general notion of reasoning (recall my argumentation in Chap. 3, Sect. 4.2; cf. Kaiser 2012).

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). The problem at this point is that 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. One could, however, 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 Sect. 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). Hence, 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. I think Sarkar would have done better to give up his demand for neutrality and to accompany other opponents of Nagel's model, as 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. I now explicate and discuss these three criteria of reductivity in more detail.

The first criterion, "fundamentalism", captures Sarkar's impression 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[23] (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* (*abstract*) *level* than the (abstract) 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).[24] A typical example of a reductive explanation that satisfies the first and second criterion (but not the third) 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

---

[23] 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.

[24] 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).

these abstract hierarchical reductions (which he also calls "genetic reductions"; 1998, 103) are the most common kind of reductive explanations.[25]

Reductive explanations that satisfy all three criteria are by far most common (in molecular biology). They are characterized by a hierarchy, which is spelled out *spatially* (thus, the third criterion is named "spatial hierarchy"[26]). In reductive explanations 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. Reductive explanations of this third kind are also called part-whole explanations (because part-whole relations are assumed to play an important role in determining levels; cf. Kaiser manuscript a; Chap. 6, Sect. 1.2). A typical example of this kind of reductive explanation 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. According to Sarkar's account this explanation is reductive since it satisfies the "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 of this kind are called "strong reduction[s]" (1998, 44) (or "physical reductions"; 1998, 136). 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.

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):

Sarkar's Account

Explanations are reductive iff

(1) P and E belong to different realms, whereas the realm of E is assigned a fundamental status (*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 (*abstract hierarchy*) or

(3) criteria (1) and (2) are fulfilled and the hierarchy referred to in (2) is spelled out spatially (*spatial hierarchy*).

---

[25] 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.

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

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 several 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. But also Sarkar's account is not fully convincing. In what follows, I will highlight three shortcomings of his account.

First, Sarkar's analysis is restricted in scope since it considers only reductive explanations in genetics and in molecular biology. Although this allows an in-depth investigation of the explanatory practice of these two disciplines, it has the drawback that it sheds light only on a small range of reductive explanations in biology. Sarkar's account thus might fail to capture the diversity of reductive explanation present in contemporary biology (recall Chap. 2, Sect. 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.[27] In doing so, he ignores part-whole explanations, in which the parts are located on a lower spatial level than the phenomenon to be explained, but not on the lowest level of molecules. As I will elaborate in Chap. 6, Sect. 2.4.3, 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 proven 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).[28] Since explanations are representations and as such always partial (or selective), any

---

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

[28] 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)

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 by distinguishing something as fundamental or as non-fundamental.[29] It thus 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 prior or as fundamental and ignoring other factors as irrelevant. Contrary to Sarkar, 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).[30] 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 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 (Sect. 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 (Sect. 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 (Sect. 2.3.3)? I conclude by revealing the objections their account encounters (Sect. 2.3.4).

---

[29] 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'.

[30] 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 and Hüttemann 2011).

2.3.1 Methodological Framework

One reason why the analysis of Hüttemann and Love is valuable to my project is that they aim at a target that is similar to mine and they make similar methodological presuppositions as I do (see Chap. 2). 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)

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 can be distinguished from non-reductive explanations. Their demand for "application" corresponds to my criterion of descriptive adequacy (Chap. 2, Sect. 1) and their requirement of "relevance" is in line with what I characterized as the potential usefulness of philosophy to biological practice (Chap. 2, Sect. 6).

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. I therefore do not shrink from calling my account a "theory" of reductive explanations.[31]

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, *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 macromolecules (e.g., DNA strands, amino acids, m- and t-RNAs, etc.). On the other hand, it comprises claims about causal processes such as the 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

---

[31] 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.

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-whole reductions" (2011, 527),[32] 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.[33] 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

---

[32] However, in the next section I point out that it remains unclear whether Hüttemann and Love identify reductive explanations with part-whole explanations or whether they assume that there exist also non-reductive part-whole explanations, as some of their remarks suggest.

[33] 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 Chap. 6).

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. They thus 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 aspect, 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. This seems to be exactly what Sarkar aims to capture by his "spatial hierarchy" criterion of reductivity (recall Sect. 2.2.3).

But caution is needed here. The way in which Hüttemann and Love specify the notion '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 which of these three readings of fundamentality they actually use. Sometimes they indicate that they refer to the third reading by using the phrase "qualified fundamental level" (2011, 529) or "qualified sense of fundamentality" (2011, 538). But 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 (third reading) or if one assumes a certain notion of level, as I will explain in detail in Sect. 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)

In this quotation Hüttemann and Love argue that 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*.[34]

---

[34] 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).

2.3.3 The Reductivity of Part-Whole Explanations in Biology

What do Hüttemann's and Love's three aspects of reductive explanations 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).[35] But this holds also for biological explanations that are non-reductive; many of them exhibit a temporal character, too. Temporality thus does not help to distinguish reductive from non-reductive explanations. 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). 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. In short, intrinsicality and fundamentality are features that determine the reductive character of biological explanations, whereas temporality is *not* such a feature. It might be that Hüttemann and Love would agree to this difference, but they fail to sufficiently emphasize it. What is more, passages such as "explanations are *reductive* 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, which is not the case.

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 claim 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 both intrinsicality and fundamentality fail.[36] Consider the examples they discuss. The first case, the failure of intrinsicality, is

---

[35] 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).

[36] 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 Sect. 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 Sect. 2.3.2).

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. But 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.[37] 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. Accordingly, 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 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 and 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 actively facilitate folding. In their view, 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 (Sect. 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*.[38] Otherwise, what counts as fundamental is identified with what is intrinsic to the

---

[37] 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$_{IF}$.

[38] Strictly speaking, they claim that only fundamentality is decoupled from intrinsicality, not vice versa.

system of interest because intrinsic factors are parts of the system and – given that levels are determined by part-whole relations – located on a lover level than the system. 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 Chap. 6). But 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 Shortcomings of Hüttemann's and Love's Account

The analysis of Hüttemann and Love provides several notable insights into the character of reductive explanations in biology and 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 Chap. 6. But their account also faces serious difficulties, which I discuss in this section. The first set of objections concerns Hüttemann's 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 explanations.

First, as I have already demonstrated in Sect. 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. We need to understand what it means for an explanatory factor not 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.[39] 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). But 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, concerning the aspect of temporality further objections arise. To begin with, Hüttemann and Love do not sufficiently highlight the different status temporality has compared to the status of intrinsicality and fundamentality. Contrary to these two aspects, temporality is not a determinant of the reductivity

---

[39] At least if the notion of a spatial level is defined locally (see Chap. 6).

of explanations (recall Sect. 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 (Sect. 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 entities and activities bring about the system's behavior. In other words, they conceive the *temporal* organization of the components of a mechanism as crucial for an explanation (e.g., Craver and 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 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 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 (from t to e.g., t^*) are cited as explanatorily relevant.[40] Consider the example of photosynthesis. In order to give a part-whole explanation of this phenomenon a temporally extended 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

---

[40] 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.

character of part-whole explanation in biological science Hüttemann and Love disregard this important kind of temporality.[41]

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.[42] According to this interpretation, only a subset of part-whole explanations would be *reductive* part-whole explanations. This, however, conflicts with Hüttemann's and Love's initial thesis that reductive explanations can be identified with 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. 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). But it seems odd to call an explanation that refers to more than just to the parts of a system a "part-whole explanation". Hence, according to this view part-whole explanations are all reductive explanations, but there may be reductive explanations that are not part-whole explanations. Table 4.1 illustrates these interrelations.

Table 4.1 The relation between intrinsicality, fundamentality, the reductive, and the part-whole character of biological explanations

Intrinsicality	Fundamentality	Reductive character	Part-whole character
-	-	-	-
-	+	+	-
+	+	+	+

---

[41] I suppose that this is why Love and Hüttemann add the bracket "(or $t_1, \ldots, t_n < t^*$)" to claims such as "the state or properties of the parts and their interactions at t (or $t_1, \ldots, t_n < t^*$) 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".

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

3 Interim Conclusion

In this chapter, I point out two different perspectives on explanatory reduction in the philosophy of biology that must be distinguished. The first perspective is the notion of explanatory reduction that implicitly underlies Rosenberg's defense of explanatory reductionism (Sect. 1). He treats explanatory reduction as a relation between an ultimate how-possible and a molecular proximate why-necessary explanation of the same phenomenon. The other perspective focuses on individual biological explanations and tries to specify the conditions under which explanations exhibit a reductive character (Sect. 2).

One central outcome of this chapter is that the second perspective yields a strategy of analyzing explanatory reduction in biology that seems most promising. This is not least because Rosenberg's perspective faces some serious objections. He fails to detach his account from the Nagelian framework and his view of 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 actually is. Furthermore, Rosenberg draws an artificial line between questions of reduction and questions of explanation that cannot be sustained.

Another goal of this chapter is to identify those ideas about explanatory reduction that seem adequate and fruitful and which I thus should take into account when developing my own approach. Four ideas are particularly notable: First, 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 and Love 2011), and lower-level explanations (Sarkar 1998; Hüttemann and 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 and to understand why this is so (see Chap. 6, Sect. 5).

Third, Sarkar (1998) argues that an appropriate analysis of reductive explanation should focus on substantive issues, not on formal ones. This is a crucial insight that stands behind much of the criticism that can be raised against Nagelian models of theory reduction (recall Chap. 3, Sect. 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 that exist in the world (more on this in Chap. 6, Sect. 6).

Fourth, Sarkar, Hüttemann, and Love identify different criteria or aspects of reductive explanations in the biological sciences (Sarkar: fundamentalism, abstract hierarchy, and spatial hierarchy; Hüttemann and Love: intrinsicality, fundamentality, and temporality). The criteria that I will suggest in Chap. 6 deviate from theirs in important respects, but my analysis also profits from notable insights that their accounts comprises. 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. Furthermore, Hüttemann and Love are the first who recognize a difference between the claim that something is located on a lower level (what they call fundamentality) and that something is located inside the spatial boundary of an object or system (what they call intrinsicality).

Chapter 5

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 2004b, 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)*

Contents

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 and 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

M.I. Kaiser, *Reductive Explanation in the Biological Sciences*, History, Philosophy and Theory of the Life Sciences 16,
DOI 10.1007/978-3-319-25310-7_5

which identifies different features that adequate or good explanations must possess) is correct. In Sect. 4 I will side with van Regenmortel and show why discussions about whether explanatory reductionism is true or not *in fact* amount to specific questions about scientific explanation.

Recall, however, that the target of my analysis is not reductionism, but reduction (Chap. 3, Sect. 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 also be of interest – but only insofar as they concern issues that impact my account of reduction. Accordingly, in this chapter I discuss only those questions about explanation 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 of 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; cf. Craver 2014)? In other words, where does the explanatory force in explanations stem from? At the outset of this chapter (Sect. 1) I give a very brief introduction to the different answers that have been proposed to this question and distinguish *covering-law* (*CL*) *models* of explanation from *causal-mechanical* (*CM*) *models* of explanation. Under CL models I subsume Hempel's deductive-nomological (DN) model, his inductive-statistical (IS) model (Hempel and Oppenheim 1948; Hempel 1965), as well as the unificationist account of explanation advocated by Michael Friedman (1974) and Kitcher (1981, 1989, 1999b). CM models have been worked out by Wesley Salmon (1984a, 1994, 1997) and further developed by the New Mechanists (e.g., Machamer et al. 2000; Glennan 2002; Bechtel and Abrahamsen 2005; Craver 2007a) and by proponents of an interventionist theory of causation (e.g., Woodward 2003; Strevens 2008). Since the debate about the nature of scientific explanation is both an extensive and a well-known debate I will only introduce some core assumptions to freshen up the readers memories and suggest some literature for looking up details. The only two issues that I will elaborate on as they play a central role in Chap. 6 are the mechanistic account of explanation (Sect. 1.2) and the question what it means when proponents of the CM model call their conception of explanation ontic (Sect. 2). The latter is important to me since I also characterize my own account of explanatory reduction as ontic and, in doing so, draw on how this term is understood in the debate about explanation.

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. 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). I point out that the truth of these assertions must be judged separately with regard to debates about the adequate account of reduction (Sect. 3) and with regard to debates about the correctness of reductionism (Sect. 4). In Sect. 3 I show why the question what constitutes the reductive character of biological explanations (the *question of reduction*) does not boil down to the question 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 virtue of my account.

In Sect. 4 I turn to the question whether and in which respect debates about explanatory reductionism depend on discussions about explanation. I begin by examining wherein these discussions about explanation consist, to which questions about reductionism allegedly amount to (Sect. 4.1). The result is that those discussions about explanation that matter to reductionism concern questions about the adequacy of higher- and lower-level explanations. In Sect. 4.2 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 (rather than by one's stance on whether CL or CM models are adequate). Hence, in Sect. 4.3 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 (Sect. 4.4).

The issues I approach in Sect. 4 occupy center stage in the long-standing debate about the truth of explanatory reductionism. But 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. I think these issues are in fact *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 (e.g., which examples of reductive explanations are analyzed), 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 Chap. 6), which needs to be taken into account. The revised version of van Fraassen’s pragmatic account of explanation that I develop in Sect. 4.3 ensures that all different kinds of explanation that can be found in actual biological practice are considered. Hence, I need the discussions presented in Sect. 4 in order to reject views of explanation that are incompatible with my account of explanatory reduction.

1 Accounts of Explanation

Philosophical theories of explanation seek to elucidate the nature of explanation. That is, they aim at identifying those features that all explanations share (i.e., the “essential characteristics” Hempel and Oppenheim 1948, 135) and that distinguishes them from mere descriptions (or, more generally, from non-explanations or inadequate explanations[1]). They offer an answer to the question what it is to achieve an understanding of *why* something happens, as opposed to merely knowing *that* something happens.[2] Only the answers they give differ. 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. 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.

---

[1] I use the term ‘explanation’ as a success term. This is why, 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. Despite this decision I sometimes use phrases such as ‘adequate explanation’ (which is, strictly speaking, done twice) since this is the easier way of expressing something clearly.

[2] 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 Sect. 4.3.3).

1.1 Covering-Law (CL) Model

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 (e.g., a statement, representation, or proposition) and not something in the world itself (e.g., Salmon 1992, 10; Craver 2007a, 27, 2014; see also Sect. 2).

According to Hempel’s *deductive-nomological* (*DN*) *model* (Hempel and 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[3] 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:

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

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

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.

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. Three major objections to the DN model should be briefly mentioned

---

[3] 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.

here: first, cases like the flagpole example reveal that some explanations exhibit asymmetric or directional features, to which the DN model is insensitive; second, another kind of example shows that the DN model fails to sort out descriptions, which include explanatorily irrelevant information, as non-explanations (e.g., deducing "John Jones fails to get pregnant" 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). Both objections show that the DN model fails to provide sufficient conditions for explanation. Third, 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.[4]

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

1.2.1 Salmon's Account

The CM model of explanation is typically traced back to Salmon (1984a, 1994, 1997). Initially Salmon attempted to characterize explanation in purely statistical terms (1971). But in 1984 he abandoned this attempt and advocated a CM view of explanation, which he called "the ontic conception" (1984b, 296; more on this in Sect. 2) 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).[5] But the key concepts of his account remained the same: causal processes (distinguished from pseudo-processes) and causal interactions.

---

[4] In response to this objection some have adopted the so called "hidden structure strategy" (Woodward 2003, 159, 2011, Section 2.6): Hempel (1965), for example, 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.

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

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 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, contrary to Salmon, the mechanistic conception of explanation focuses on constitutive causal explanations.[6] Salmon also uses the notion of a causal mechanism to characterize his ontic conception of explanation.[7] However, 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 identical to causal explanations.

Salmon's CM model encounters serious objections, which I can only point at here. Most importantly, it has been argued that Salmon's CM model fails to specify which features of a causal process are explanatorily relevant to the phenomenon to be explained. For example, we have the intuition that John Jones taking birth control pills, contrary to his gender, is causally and explanatorily irrelevant to his childlessness. But Salmon's CM model overlooks this difference in relevance because when John Jones ingests birth control pills there are also spatio-temporally continuous causal processes and causal interactions at work (e.g., the pills dissolves, certain

---

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

[7] For instance, he writes that "explanatory knowledge is knowledge of causal mechanisms… that produce the phenomena with which we are concerned" (Salmon 1989, 128).

constituents are transported into the blood, etc.; cf. Hitchcock 1995). Another objection that applies particularly to Salmon's actual account is that it regards causal relations on higher levels to be dependent on the causal connections between their fundamental physical constituents. It can be argued that this reductionist strategy "leads us away from the right level of description" (Woodward 2011, Section 4.3).[8]

1.2.2 The Mechanistic Account

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 and Abrahamsen 2005; Craver 2007a; Craver and Darden 2013). 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 share the *basic idea* that explanations situate an explanandum 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 are adequately represented is of interest to whether the representation is explanatory or not.

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. CM models do 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.[9] 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 (cf. Kaiser and Craver 2013).

In what follows I briefly introduce the mechanistic conception of explanation since this version of the CM model is most relevant to my analysis of reductive explanation (the relation between mechanistic and reductive explanations will be specified in Chap. 6, Sect. 5). The first thing to note is that there exists an important difference between Salmon's CM model and the mechanistic account. Salmon recognizes two aspects of causal explanations, the etiological and the constitutive

---

[8] 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.

[9] 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).

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 due to the fact that examples of this kind occupy center stage in debates about causation: 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.[10] By contrast, the mechanistic account focuses on *constitutive* causal explanations, that is, on explanations that describe the underlying causal processes and interactions (in mechanistic terms: the underlying causal mechanism) that constitute the phenomenon to be explained. The New Mechanists also discuss etiological mechanistic explanations, but to a much lesser extent (Kaiser and Krickel forthcoming).

According to the mechanistic conception of explanation, the behavior of a system (i.e., the explanandum phenomenon) is constitutively mechanistically explained by describing how certain entities and activities are organized together such that they bring about the behavior to be explained (e.g., Craver 2007a, 2–9, 121–162). Three further theses about mechanistic explanations are relevant for my purposes: First, constitutive mechanistic explanations are said to be a special kind of *part-whole explanations* because the behavior of a system (the whole) is explained in terms of the mechanistic components, that is, in terms of entities and activities that constitute the phenomenon in question. Even though this is correct, it is important to notice that "being a part of the phenomenon to be explained" need not fall together with "being a component of the underlying mechanism". The mechanism of muscle contraction includes also molecules (e.g., neurotransmitters) that are not parts of the contracting muscle fiber (the phenomenon to be explained). Moreover, the muscle fiber has several parts (e.g., chromosomes or centrosomes) that are not components of the mechanism for muscle contraction because they are irrelevant to the phenomenon in question (more on this in Kaiser and Krickel forthcoming and in Chap. 6, Sect. 5).

This leads us to the second point, the question of explanatory relevance. The New Mechanists state that in constitutive mechanistic explanations a certain behavior of a 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

---

[10] 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. Hence, the way that Salmon and others *in fact* discuss these 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 also in an etiological-constitutive manner.

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, b) to distinguish those parts of the system (together with the activities they perform), which are components of the mechanism in question (i.e., which are constitutively relevant), from other, explanatorily irrelevant factors.

Third, proponents of the mechanistic account emphasize that entities and activities that constitute a mechanism must be *spatially and temporally organized* in a specific way (e.g., Craver 2007a, 134–139; Craver and Darden 2013, 20). 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 that constitute 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 to be explained 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. First, 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 and propose a different characterization (e.g., Illari and Williamson 2012). One might object that the mechanistic approach rests on problematic concepts and metaphysical assumptions, such as the mechanists' notion of productivity and their thesis that activities make up a distinct ontological kind (e.g., Woodward 2002; Tabery 2004; Torres 2008). One might point out that the mutual manipulability criterion that is part of Craver's constitutive relevance approach is problematic (e.g., Leuridan 2012; Baumgartner and Gebharter 2015; Kaiser and Krickel forthcoming). All of these issues are very interesting but they lie outside of the scope of this book. Second, one can question the endeavor of the New Mechanists as a whole. Since their account focuses on a particular kind of explanation, namely (constitutive) causal explanations (in the life sciences), it is questionable whether they offer a real competitor to the CL model. It seems as if a real competitor must share the universalistic aspiration of the CL model, that is, it must claim to hold for scientific explanations in general.[11] On the contrary, if the mechanistic account is

---

[11] In this spirit several mechanists try to broaden the scope of the mechanistic account. For example, Craver argues that etiological and functional explanations are subtypes of mechanistic explanation (2007a, 107, 2013), Skipper and Millstein (2005) apply the mechanistic conception to natural selection explanations, and Glennan (2010) claims that it also holds for historical explanations.

thought to be a universal model of explanation that applies to *all* scientific explanations it is vulnerable to counterexamples since it does not capture non-mechanistic or non-causal kinds of scientific explanation, such as mathematical explanations, topological explanations, "aggregate explanations", or "morphological explanations" (Craver 2007a, 162).

2 What Makes a Conception 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, 2014; Glennan 2002, 343; Strevens 2008, 6f, 43).[12] Early proponents of the ontic conception of explanation were Salmon (1977, 1984a, 1989), Railton (1981), and Coffa (1974). But what exactly does the term 'ontic' mean? What is it that makes a philosophical conception of explanation ontic? Let us start by examining how Salmon employs this term.

Salmon distinguishes three basic conceptions of scientific explanation: first, *epistemic* conceptions, which treat explanations as arguments that render the event-to-be-explained expectable in virtue of the explanatory facts[13]; 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 (for further details see Salmon 1984a, 15–20, 84–134). Those philosophers who adopt the ontic conception generally regard the pattern into which events are to be fitted as a causal pattern. Accordingly, 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, a strong and a weak reading. According to the *strong reading*, what makes an account of explanation ontic is the thesis that explanations are entities (e.g., objects, facts, or relations) that *exist in the world*, independently from scientists discovering them, from inquirers requesting them, or from speakers

---

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

[13] Salmon characterizes inferential conceptions, such as Hempel's DN and his IS model, and erotetic conceptions, such as van Fraassen's pragmatic account of explanation (which will be discussed in Sect. 4.3), as *epistemic* conceptions of explanation.

uttering them. In the framework of the CM model, these explanatory entities in the world are assumed to be causes or parts of the causal structure of the world. This strong interpretation of 'ontic' can be found most clearly in the work of Craver (2007a, 27, 33, 200, 2014) and Strevens (2008, 6f, 43). For example, Strevens claims that explanations are "something out in the world, a set of [causal] facts to be discovered" (2008, 6). Craver makes it even more explicit: "The explanations are in the world" (2014, 50). In more detail:

> [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. (Craver 2007a, 27; my emphasis)

Craver and Strevens recognize that the term 'explanation' is often used to refer to epistemic units, such as descriptions, representations, models, or explanatory texts that convey information or communicate scientific knowledge (e.g., Craver writes: "sometimes explanations are texts" 2007a, 27). They emphasize, however, that one should give precedence to the ontological sense of explanation and that it is important to shift the attention away from the representations used in explaining phenomena and toward the causal structure of the world. Only "ontic explanations" (Craver 2014, 40) or "objective explanations", as they argue, provide "systematic clues about the nature of explanation itself" (Strevens 2008, 6) and, thus, constitute "the correct starting point" (Craver 2007a, 27) for developing an account of explanation. Among the critics of the ontic conception, this strong reading is endorsed for instance by Wright (2012) and it is what Wright and Bechtel refer to as understanding the ontic conception "literally" (2007, 49).

According to the alternative, *weak reading* of the ontic conception, explanations are not mind-independent things in the world itself but *epistemic entities* such as descriptions, representations, propositions, or explanatory texts. 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 that, for instance, 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)[14]

But if it is not the thesis that explanations are things in the world, what is it then that makes a conception of explanation ontic and distinguishes it from epistemic

---

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

conceptions? 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. The explanatory force thus arises not from logical features of statements or from relations of nomic expectability. Instead, what makes a description explanatory, according to the weak reading of the ontic conception, is that it truly represents certain features of a mind-independent reality, namely the causal structure of the world.

In a nutshell, these two views of what the ontic conception of explanation amounts to draw the distinction between ontic and epistemic accounts differently. According to the strong reading, ontic and epistemic accounts primarily differ in how they answer the question '*What kind of entity is an explanation*?'. Ontic accounts regard explanations as entities in the world, whereas epistemic accounts regard explanations as representations or descriptions of these entities in the world. By contrast, according to the weak reading, both ontic and epistemic accounts agree that explanations are epistemic entities but differ in how they answer the question '*What determines the explanatory force of an explanation*?'. Hence, ontic accounts claim that whether a given representation has explanatory force depends on whether it truly represents certain entities in the world – not on whether the explanation consists of statements between which certain logical relations hold, as the epistemic account claims. In other words, according to the weak reading, 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, so the proponent of the weak reading of the ontic conception, 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. John 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* (1998, 77; my emphasis).

Another way to express the weak reading of the ontic conception of explanation is to say that ontic and epistemic accounts identify different "normative constraints on explanation" (Illari 2013, 241), that is, different standards for what makes an explanation a good explanation. Ontic conceptions of explanation emphasize "ontic constraints" (Illari 2013, 242), which are worldly features such as the components of a mechanism that is responsible for the phenomenon to be explained. Epistemic conceptions, on the other hand, regard "epistemic constraints" (Illari 2013, 245) as more fundamental. Epistemic constraints are, for instance, mathematical techniques or psychological features of those who seek explanations.[15]

---

[15] In her paper, Phyllis Illari argues that philosophers should not focus on either kind of constraint, but rather recognize both and seek an integration of them.

My main claim in this section is that both the strong and the weak reading are *equally legitimate* views of what characterizes the ontic conception of explanation and distinguishes it from an epistemic conception of explanation. Both readings are plausible interpretations of Salmon's claims. Even more, Salmon himself recognizes and accepts these two different readings of the ontic conception:

> Proponents of this conception [i.e., the ontic conception] can speak in either of two ways about the relationship between explanations and the world. First, one can say that explanations exist in the world. The explanation of some fact is whatever produced or brought it about. […] Second, the advocate of the ontic conception can say that an explanation is something – consisting of sentences or propositions – that reports such facts. It seems to me that *either way of putting the ontic conception is acceptable*. (Salmon 1989, 86; my emphasis; see also 1992, 35)

Furthermore, both readings are coherent positions, which do not result in a conflation of ontic and epistemic conceptions. As I pointed out, both readings retain the distinction but draw the line between ontic and epistemic conceptions differently. Finally, both readings of the ontic conception of explanation can be found in the literature. Among the proponents of the mechanistic conception of explanation, Craver (2007a, 2014) for instance, defends the ontic conception of explanation in its strong reading, whereas Glennan (2002) endorses the weak reading of the ontic conception.

One might still try to argue that the weak reading identifies conceptions of explanation that are not ontic in the proper sense. This might be an argument from authority (or popularity): Those who prominently defend the ontic conception of explanation, such as Craver and Strevens, understand it according to the strong reading. Thus, we should use the term 'ontic' in the very same way. I agree that using the term 'ontic' in a different, less popular way (namely according to the weak reading) might provoke misunderstandings. This is why I spend so much effort on distinguishing the two readings of the ontic conception. And this is why I emphasize that when I characterize my account of reduction as ontic, I refer to how the term 'ontic' is understood in the weak reading of the ontic conception of explanation (see Chap. 6, Sect. 6). The argument from authority underlines the necessity to be very clear in how one understands the term 'ontic' (for this purpose one might, for instance, distinguish between the two terms '$\text{ontic}_{\text{strong}}$' and '$\text{ontic}_{\text{weak}}$'). But this argument achieves no more: It does not show that the more popular strong reading of the ontic conception yields the only legitimate or proper understanding of the term 'ontic'. Other versions of this objection seem to boil down to a mere dispute about how to use words, which is also not convincing at this point.

If one accepts my argument that the strong and the weak reading of the ontic conception are equally legitimate, one might still wonder why I use the term 'ontic' according to the weak reading. This decision is because, in my view, the ontic conception of explanation understood in the strong sense faces serious problems. 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 at least three respects. First, regarding explanations as things in the world completely decouples explanation from understanding. The causes that lead to or make up a certain phenomenon are an explanation of this phenomenon regardless of whether the tracing of the causes renders this phenomenon intelligible and promotes its understanding. Whether we understand the phenomenon better when we know what its causes are is simply irrelevant to the causes being explanatory. I think 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, such as 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 do not), this acceptance does not imply that they are the *only* adequate target of philosophical analysis. Neither Craver nor Strevens offer a real argument for why philosophers should be concerned only with explanations as things in the world.[16]

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 (more on this in Sect. 4.3).

To conclude, I introduce two different views of what the ontic conception of explanation amounts to. The strong reading interprets the ontic-epistemic dispute as a dispute about what explanations are – entities in the world or representations of these entities. By contrast, the weak reading locates the disagreement between the ontic and the epistemic account in differing views about what determines whether a representation has explanatory power – true representations of entities in the world or logical relations between sentences. I argue that both of these readings are legitimate: Salmon accepted them, they are coherent views that do not conflate the ontic-epistemic distinction, and they can be found in contemporary literature. What speaks against the strong reading is that the ontic conception, so understood, encounters serious objections. On these grounds I will understand the term ‘ontic’ according to the weak reading when I characterize my own account of explanatory reduction as ontic (more on this in Chap. 6, Sect. 6).

---

[16] Strevens, for example, 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 precedence to the ontological sense of explanation. 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 2013).

3 Explanation and Explanatory Reduction

Recall the quotations from which this chapter started. Some authors claim that the debate about reduction and reductionism is, basically, a debate about what constitutes a good or adequate explanation (van Regenmortel 2004b, 145). Others explicitly deny this. They state that questions about explanatory reduction(ism) cannot be decided by settling for a certain position in the debate about explanation (Rosenberg 2006, 41). So, which of them is right?

The first point to note is that this question needs to be answered *separately* with respect to reduction and with respect to reductionism. In a nutshell, the question what explanatory reduction in biological practice is, that is, what constitutes the reductive character of biological explanations, does *not* boil down to the question of what characterizes an adequate explanation. 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 next section). In contrast, the question of whether explanatory reductionism is true or not in fact amounts to particular questions about explanations. With regard to reductionism (but not with regard to reduction) I thus side with van Regenmortel and oppose Rosenberg. Before I specify which questions about explanation are relevant to discussions about explanatory reductionism (Sect. 4) I dwell on the relation between my account of explanatory reduction and questions about explanation (this section).

3.1 Different Questions

As I introduced in Sect. 1, covering law (CL) models and causal mechanical (CM) models differ in what they regard as the nature of explanation. They provide different answers to the question what constitutes the explanatory force of explanations and distinguishes explanations 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. 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; recall Sect. 2) 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). 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* 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.[17] These additional criteria will be specified in the subsequent chapter. In sum, the difference between the two questions is the following:

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?

But 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 any *interconnections* between them. 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 (Chap. 3, Sect. 4) one might contest that 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, although it has its shortcomings, too. 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

---

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

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. Hence, my account of explanatory reduction can remain neutral regarding the conflict about the nature of explanation.

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 to explicitly argue in favor of the 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 advisable 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 Sect. 2 I 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 unconvincing. 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 Sect. 4). It takes for granted that explanations are epistemic units (such as representations or descriptions), not things in the world. But this decision excludes only more radical versions of the CM model (like Craver's and Strevens'). It leaves the question of explanation still unanswered.

3.2 Some Matters of Terminology

I just argued that 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 concept of explanatorily relevant "factors".[18] This terminology has two merits: first, it leaves open whether these factors are causal. Since most reductive explanations in the biological sciences exhibit a temporal character (Hüttemann and Love 2011; Love and 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-

---

[18] I use the term 'factor' as an ontological term (i.e., 'factor' refers to some entity in the world). At the same time the term should be as ontologically uncommitted as possible. It may refer to causes, relations, processes, regularities, universals, facts, activities, or whatever one's ontology and theory of explanation demand. Sets of factors thus are what is being represented in explanations, not what constitutes the explanation itself.

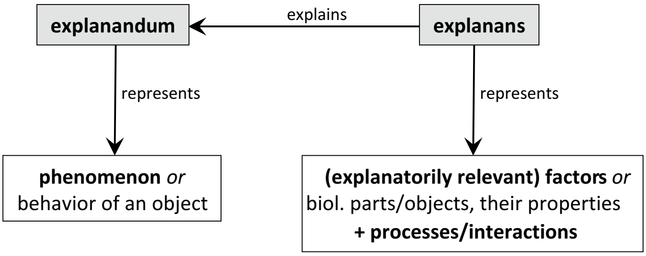

Fig. 5.1 Parts of an explanation and what they refer to

causal explanations in biology, too (e.g., gene-selectionist explanations).[19] My way of speaking about "factors" that are referred to in the explanans leaves room for this kind of reductive explanation.[20] Second, the notion of explanatorily relevant factors is uncommitted with regard to the question whether the explanans contains generalizations or law statements. Advocates of the DN model will insist that among the explanatorily relevant factors there must be laws. Whereas 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.

I sometimes deviate from speaking about "factors" that are described in the explanans. In addition, I also state that a phenomenon or a behavior of a biological object or system is explained by appealing to its biological parts or by reference to certain objects, their properties, and the processes and interactions they are engaged in. Figure 5.1 illustrates this.

This terminology is common in the philosophy of biology, in particular, in debates about reduction, part-whole explanation, and mechanistic explanation. But 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, 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.

---

[19] Gene-selectionist explanations are evolutionary explanations that explain evolutionary processes exclusively by appealing to selection processes at the level of genes (see Dawkins 1976; Wimsatt 1980). But it remains a controversially disputed question whether natural selection explanations are to be characterized as non-causal explanations (e.g., Sarkar 2005, 117–143).

[20] 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).

4 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 (Sect. 4.1). After having clarified this I argue that the dispute about explanatory reductionism in fact boils down to answering specific questions about explanation, namely questions about the adequacy of higher- and lower-level explanations (Sect. 4.2). 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 Sect. 4.3 I examine van Fraassen's pragmatic theory of explanation as well as the objections Rosenberg raises against so-called 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 (Sect. 4.4).

Before I start, let me address a possible objection. Having in mind that I stressed that my goal in this book is to develop an account of reduction, not to dispute the truth of reductionism (see Chap. 3, Sect. 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 project? I think this is not the case. The issues I discuss in this section are *relevant* to the project of seeking an adequate account of explanatory reduction, even though they are only indirectly relevant. Certain views of explanation, such as Rosenberg's, must be proven to be incorrect since their correctness would restrict the set of "empirical data", on which my analysis is built (e.g., examples of reductive explanations), 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, 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 and therefore must be rejected.

4.1 Which Questions About Explanation Are Relevant to Reductionism?

Consider the arguments van Regenmortel provides. 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. (2004b, 145)

This statement reveals that van Regenmortel is primarily concerned with *in-practice* reductionism (recall Chap. 3, Sect. 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 is plausible to say 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[21] (Byerly 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.

4.1.1 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 the alternative, 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 as being a "terrible explanation" 1975, 296).

In a similar vein, Kitcher claims that the general phenomenon that genes on nonhomologous 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.

---

[21] 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. But as such a way of speaking is less intelligible I accept the minor inaccuracy involved in using the term 'inadequate explanation'.

Putnam's and Kitcher's arguments have been critically discussed by various authors and from different perspectives (reductionist and antireductionist).[22] What is common to all these discussions is that they concern the question of what the *right level of explanation* is. An explanatory antireductionist (as Kitcher) argues that a higher-level phenomenon (e.g., 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 (as 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).

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). But 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

---

[22] To name only some philosophers of biology: Waters 1990, 131–134, 2008, 244–249; Schaffner 1993, 478–481; Sober 1999; Rosenberg 2006, 32–47.

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 Chap. 4, Sect. 1.1).

4.1.2 Relevant Questions About Explanation

As I will elaborate in the next section, 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, as 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.

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 (or, in imprecise terms, from inadequate explanations).

Second, not all discussions about the adequacy of explanations are relevant to the dispute about explanatory reductionism. Only those that concern *levels of explanation* are crucial. 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.[23] 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

---

[23] For the corresponding notion of a level of organization see Chap. 6, Sect. 1.2 and Kaiser (manuscript a).

explanation. In contrast, reductionists typically defend the explanatory priority of lower levels or the lowest level of explanation (as Rosenberg).[24]

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*, such as explanatory interests, background information, etc., might play in determining what counts as explanatory. It is crucial to note that the debate about the pragmatics of explanation is *much more relevant* to the topic of reductionism than the 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.[25] 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.

There is, however, an important restriction of this third thesis, which needs to be emphasized. 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 (recall Sect. 1.2) 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 components, which – according to the mechanistic account of levels (see Chap. 6, Sect. 1.2) – can be said to be located on a *lower level* than the phenomenon to be explained. Thus, the mechanistic conception focuses on lower-level

---

[24] 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.

[25] 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. But if one abandons Nagel's understanding of reduction, the dispute between CL and CM models of explanation loses significance, too.

explanations.[26] Most mechanists confess that their account is limited in scope (i.e., that it does not capture all kinds of explanation), which is why they can allow for explanations that are not 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) 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 that are relevant to explanatory 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). 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, pragmatic dimensions of explanation will be examined in more detail in Sect. 4.3.

4.2 Why the Reductionism Dispute 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

---

[26] In Chap. 6, Sect. 5 I will, however, argue that not all mechanistic explanations need to be lower-level explanations. Some mechanistic explanations include factors that are external to the corresponding system and that cannot be said to be lower-level factors (at least not according to my account of levels of organization; cf. Kaiser manuscript a).

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 strongly supports this view (2006, 32–55, 178–181).

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). But the reasons Rosenberg adduces 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 (i.e., lower level) 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. The antireductionist will, however, 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. But 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 next section). By contrast, a reductionist will reject this thesis and claim that only the molecular parts of the ideal explanatory text constitute adequate explanations. This means that contrary to Rosenberg's view 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.

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. In the previous section I have shown that the two questions can*not* 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.

4.3 Pragmatic Dimensions of Explanation

4.3.1 Van Fraassen's Pragmatic Theory of Explanation

The most prominent pragmatic account of explanation was developed by Bas van Fraassen (1977, 1980).[27] 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 example, we ask 'Why is the muscle fiber contracted (rather than relaxed)?' or 'Why is the muscle fiber (rather than the spinal nerve) contracted?'. 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" (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 $\langle P_k, X, R\rangle$.[28]

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, also the adequacy of an explanation depends on the context. 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 (such as CL 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

---

[27] Another important representative of the erotetic version of the epistemic conception of explanation (to adopt Salmon's terminology) is Peter Achinstein (1983).

[28] For more details on van Fraassen's theory of explanation see 1980, 134–157. For a summary of van Fraassen's account see Kitcher and Salmon 1987, 317–319.

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.

But 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 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?

4.3.2 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 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., which phenomenon one seeks to explain) may be affected by the research interests and the background knowledge in a certain scientific field, this does not imply that also the explanatory relevance of certain factors is determined by pragmatic factors. 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; recall Sect. 2), 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 also 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. 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 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 I think 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 sketch what I think is a promising intermediate way of refining van Fraassen's account.

4.3.3 Types of Explanation and Genuine Relevance Relations

Let us start by recalling some of the methodological preliminaries that I discussed in the first chapter of this book. In Chap. 2, Sect. 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 from biological practice were not of particular interest for understanding 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 biological practice. In accordance with what I said about balancing specificity against generality (Chap. 2, Sect. 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 $R_1,\ldots R_n$, 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, as constitutive 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 or even as mechanistic explanations (Glennan 2010). Philosophy of biology is far from having established a complete list of types of biological explanations (see also Brigandt 2013). Moreover, 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, b). Hence, the philosophical task of identifying and characterizing those relevance relations that belong to the set of *genuine* (or "objective" Salmon 1984a, 131) relevance relations $R_1,\ldots R_n$ has to be left to future philosophical work.

But in order to reject Rosenberg's subjectivism-reproach we are not reliant on having found 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 my 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 relations $R_1,\ldots R_n$.[29] 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., $2\ H_2O + 2\ NADP^+ + 3\ ADP + 3\ P_i + \text{light} \rightarrow 2\ NADPH + 2\ H^+ + 3\ ATP + O_2$). In other words, physiologists seek compositional causal explanations (which invokes relevance relation R_2, e.g., Craver's "constitutive relevance"), not covering-law explanations (which invokes a different relevance relation R_5).

But even the choice of which of the genuine relevance relations $R_1,\ldots R_n$ applies to a certain explanandum phenomenon is not affected by *subjective* pragmatic factors, that is 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 (i.e., is "subjective"), or that it is arbitrary.

---

[29] 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).

In addition, the relevance relations 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. But 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.[30] 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.

4.3.4 How Pluralistic Is This Account?

The view of explanation I just presented is closely connected to what has been called "pluralism about the character of scientific explanation" (Brigandt 2013, Section 2) elsewhere. Three different theses of *explanatory pluralism* must be distinguished. The refined version of Van Fraassen's account is committed only to the former two.

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 R 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.[31]

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 R. 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; see Chap. 2, Sect. 4.3). Other philosophers reject this latter kind of

---

[30] If this were the case we would say that these different specifications constitute distinct relevance relations.

[31] 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.

explanatory pluralism. They insist that all adequate explanations of a certain phenomenon can be integrated into something as 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. But 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 for example the case studies presented by Kellert et al. (2006) the thought arises that Railton's picture might be "too neat 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 and 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 and explanatory interests of individual 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 $R_1,\ldots R_n$ 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.

4.4 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 $R_1,\ldots R_n$ (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. But an investigation of these different theories of causation goes beyond the scope of this book. 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 to 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 constitutive relevance (2007a, 139–160, b) treats only those factors as explanatorily relevant that contribute to producing the phenomenon to be explained. Mitochondria are constitutively irrelevant (i.e., explanatorily irrelevant) because they make no direct contribution to bringing about the synthesis of proteins, in other words, because the energy production of mitochondria and the protein synthesis of cells is not "mutually manipulable" (Craver 2007a, 153). 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 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, b), their activity (Waters 2007), or their substitutability (Sarkar 2005).[32] 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. But 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.

Reconsider now the results of Sect. 4.1. There I argued that discussions about explanation that matter to reductionism concern questions about the adequacy of higher- and lower-level explanations. Antireductionists, as Kitcher, state that higher level factors are explanatorily relevant to higher level phenomena, but lower level factors are not.[33] 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 developed in the previous section 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*[34] 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)

---

[32] 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.

[33] Some pragmatist-minded antireductionists endorse the more moderate claim that there is not "the right" level of explanation independent from any context.

[34] Note that this is a different notion of depth than the one Hitchcock and Woodward (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).

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.[35] However, Potochnik's 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.

5 Interim Conclusion

The purpose of this chapter is 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 also of interest but only insofar as they affect my analysis of reductive explanation.

The goal of Sect. 1 was twofold. On the one hand, I briefly introduced different accounts of explanation and thereby provided the basis, on which in Sects. 3 and 4 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 (Sect. 2). This is relevant to my work because I characterize my own account of explanatory reduction in biology also as ontic. I distinguish two different, equally legitimate views of what makes an account of explanation ontic. The strong reading interprets the ontic-epistemic dispute as a dispute about what explanations are: entities in the world or representations of these entities. By contrast, according to the weak reading, the ontic and the epistemic account differ in their views about what determines whether a representation has explanatory power: 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. When I characterize my own account of explanatory

---

[35] This is why Potochnik argues that they exhibit "different types of generality" (2009, 64).

reduction as ontic I understand the term ‘ontic’ according to the weak reading of the ontic conception of explanation (see Chap. 6, Sect. 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 (Sects. 3 and 4). 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 and with regard to debates about reductionism. In the former case the answer is ‘yes’, whereas in the latter case the answer is ‘no’.

The question of explanatory reduction does not boil down to the question of explanation: what makes an explanation reductive is different from what makes a description explanatory. Correspondingly, my account of explanatory reduction remains uncommitted with respect to whether the CL or the CM model adequately captures the nature of biological explanation. By contrast, the debate about explanatory reductionism in fact amounts to a dispute about explanation, namely to a dispute about the adequacy of higher- and lower-level explanations (i.e., whether certain levels are explanatorily prior to others). With whom one sides in this dispute is mainly affected by one’s stance on the pragmatics of explanation (rather than by one’s stance on whether CL or CM models are adequate).

To defend van Fraassen’s pragmatic theory of explanation 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 develop a refined version of van Fraassen’s pragmatic account. According to this 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. But 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 $R_1,\ldots R_n$ at all count as genuine, nor do they influence how these relevance relations are to be specified. The revised version of van Fraassen’s pragmatic account of explanation also ensures that the empirical basis on which my account of explanatory reduction is built includes the variety of all relevant kinds of explanation that can be found in actual biological practice.

Chapter 6
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 and 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)

Contents

M.I. Kaiser, *Reductive Explanation in the Biological Sciences*, History, Philosophy and Theory of the Life Sciences 16,
DOI 10.1007/978-3-319-25310-7_6

The preceding chapters clarified the main goal of this book. 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? Investigating these questions is valuable since the concept of a reductive explanation has not received sufficient philosophical attention so far and since a clear understanding of what reductive explanations are and wherein their merits and limitations lie can, in itself, potentially enhance disputes about explanatory reductionism in philosophy and in the biological sciences.

Before I can start with my analysis of what makes biological explanations reductive (Sects. 2, 3, 4, 5, and 6), some preliminary assumptions of this analysis must be disclosed (Sect. 1). The account of explanatory reduction that I develop in this book relies on certain concepts that are themselves in need of clarification. The core claim of my account will be that reductive explanations in biology exhibit three major characteristics: first, they refer exclusively to factors that are located on a lower level of organization than the biological object *Y* whose behavior is to be explained; second, they focus on factors that are internal to (i.e., that are parts of) the object *Y* and ignore or simplify environmental factors; and third, they refer to the parts of *Y* only as parts in isolation.

But what, exactly, is a level of organization? What determines whether something is a part of a biological object *Y* (the whole) and in this sense is internal to it? How are part-whole relations connected to levels, for instance, does the former fully determine the latter? The concept of a *biological part* and, relatedly, the concept of a part-whole relation in the biological realm, the concept of a *level of organization*, and the connections between these two concepts have gained some philosophical attention, but they have by far not been sufficiently specified yet (for biological parthood see e.g., McShea 2000; Winther 2006, 2011; Mellor 2008; Jansen and Schulz 2014; for the notion of a level of organization see e.g., Guttman 1976; Love 2012b; Potochnik and McGill 2012; Eronen 2015). An exhaustive analysis of what biological parts and levels of organization are and how they are related, however, lies beyond the reach of this book. I pursue these projects in different papers (cf. Kaiser forthcoming a, b, manuscript a, b) and, in this book, only introduce those claims that are essential to understanding my analysis of reductive explanation in biology.

The structure of this preliminary section is as follows: I begin with a brief overview of my account of biological parthood (Sect. 1.1). Then I present the main ideas of how the notion of biological part-whole relations can be used to develop an account of levels of organization that satisfies the constraints of an analysis of reduction in biology (Sect. 1.2). Drawing on some of the major claims I make in Chap. 2, I finally point out how I will proceed in developing my account of explanatory reduction and which methodological presuppositions underlie my analysis (Sect. 1.3).

These preliminaries clarify the conceptual and methodological ground of my analysis of reductive explanations in the biological sciences (Sects. 2, 3, 4, 5, and 6).

1 Preliminaries

1.1 My Account of Biological Parthood

An account of biological parthood explicates what it means for some object *X* to be a *biological part* of another object *Y* (the whole). The question under which conditions is *X* a biological part of *Y* is a special version of what van Inwagen has called the "Special Composition Question" (1990, 20). It is a special version because answering it means to provide an account of *biological* parthood in particular, not of parthood in general.[1] In other words, the goal is to understand part-whole relations that exist in the living world and that are studied by the biological sciences. This focus on *biological* parthood corresponds to the focus of this book: its goal is to understand reductive explanations in the *life sciences*, not in science in general, nor in non-scientific areas such as metaphysics.

The question what it means for an object *X* to be a biological part of some other object *Y* is highly relevant to my account of explanatory reduction because I claim that reductive explanations in biology focus on factors that are *internal* to the biological object *Y* whose behavior is to be explained and ignore or simplify *external* factors (more on this in Sect. 3). What is internal to the biological object *Y* in question and what is external to it (i.e., what belongs to *Y*'s context or environment[2]) depends on what is a biological part of *Y* and what is not: biological parts of *Y* are internal to *Y* and everything that is not a biological part of *Y* is external to it. In the following, I will sketch the core ideas of my view of biological parthood that underlie my account of reductive explanation. This overview will be rather sketchy, for more details see Kaiser (forthcoming a).

The starting point of my account of biological parthood is the idea that parts are *spatiotemporal* parts. Even though the notion of a spatiotemporal part is widespread, it is rarely specified. The rough idea is that parthood requires that the part in some sense is spatially as well as temporally included in the whole. But it is far from obvious what these requirements of spatial and temporal inclusion amount to. In my analysis of biological part-whole relations, I show how the notions of spatial and temporal inclusion can plausibly be specified (Sects. 1.1.1 and 1.1.2), and I argue

---

[1] Part-whole relations in general are analyzed in mereology (e.g., Simons 1987).

[2] In this book, I employ a very broad notion of environment, similar to Brandon's "external environment" (1990, 47–49) but which is not restricted to the environment of organisms (since the category of biological objects comprises much more than just organisms).

they must be supplemented by a third criterion, the condition of relevance (Sect. 1.1.3), to get a sufficient account of biological parthood.

1.1.1 Spatial Inclusion

Under which conditions is an object *X* a biological part of another object *Y*? An intuitive first answer is that *X* must be spatially included in *Y*. Chloroplasts are parts of cells because they are located inside the cell. My liver is a part of me because it is spatially contained in me. In Kaiser (forthcoming a) the requirement of spatial inclusion is specified as follows:

Spatial Inclusion
Object *X* is a biological part of object *Y* if

(1) if *Y* has a natural boundary then *X* is *spatially located inside* or *in* the region of the natural boundary.

Three specifics of how the spatial inclusion condition is formulated are crucial: first, spatial inclusion is *not sufficient* for biological parthood, other criteria must be added (which is indicated by the 'if'); second, to *avoid circularity*, the spatial inclusion condition must refer to the spatial boundary of *Y* as something that is identified independently of identifying *Y*'s biological parts (which is ensured by the notion of a natural boundary); third, spatial inclusion is *not even necessary* for biological parthood (which is why it is formulated conditionally). Let me explain these three points in more detail.

First, several examples from biological practice reveal that spatial inclusion is not a sufficient criterion for biological parthood. In the case of lichen, green algae are spatially located inside the fungus, but the algae are typically not regarded as parts of the fungus (they are regarded as parts of the lichen but that is a different object). If a doctor leaves a cotton ball inside my abdomen during surgery, we would not say that the cotton ball became a part of me just because it is spatially included in me. Finally, not any arbitrary DNA sequence of the human genome (which is spatially included in the genome) is conceived of as a real part of the genome. Only those DNA sequences that perform a biochemical activity of a certain kind or that possess a certain function (e.g., encoding proteins or binding transcription factors) are conceived of as "genomic parts" (Kaiser manuscript b). These and similar examples reveal that mere spatial containment is too permissive to exclusively determine biological parthood.[3]

Second, if the spatial boundary of the biological object *Y* (the whole) was identified simply by identifying all of *Y*'s biological parts and drawing an outer "line"[4]

---

[3] Craver makes a similar point with respect to levels (2007a, 187f).

[4] I use the term 'line' metaphorically. Since objects are extended in at least a three-dimensional space the boundary of *Y* is, of course, a surface, not a line.

around them, the condition of spatial inclusion would become circular: then *X* would be a part of *Y* if *X* was located inside *Y*, which would mean that *X* was a part of *Y*. To avoid this circularity, the demarcation of *Y* from its environment must be independent of the identification of *Y*'s parts. The notion of a natural boundary, which is a central element of my account of biological parthood, provides us with such an independent demarcation. I will just briefly introduce this concept here (for an elaboration see Kaiser forthcoming a).

The main idea behind the notion of a natural boundary is that many biological objects are surrounded by boundaries that exist in nature itself, rather than ones created by human demarcation (which is why I refer to them as natural). In other words, natural boundaries are "fiat boundaries", not "bona-fide boundaries" (Smith and Varzi 2000). They bind together the parts of biological objects and allow for distinguishing an inside (what belongs to an object Y, what is part of Y, what is internal to Y) from an outside (what belongs to Y's environment, what is external to Y). The most common paradigmatic examples of natural boundaries are the cell membrane, the blood-brain barrier of humans, the skin of mammals, the alveolar-capillary membrane that encloses the human circulatory system, the cell wall of plant cells, and the exoskeleton of insects. Natural boundaries of biological objects are commonly identified functionally as well as physically-structurally: they function as selective barriers, that is, they selectively hold together the objects that are located inside them and separate these objects (and the interactions between them) from what is outside. What is more, natural boundaries usually involve physical discontinuities or qualitative heterogeneities (e.g., of material constitution, texture, etc.). Despite the importance of natural boundaries in the biological realm, it should not be overlooked that *not all* biological objects possess such a natural boundary. This leads us to the last point.

Third, spatial inclusion (specified by reference to a natural boundary) is not necessary for biological parthood as there are biological objects that do not have a natural boundary. Maureen A. O'Malley and Dupré (2007; see also O'Malley 2014; Kaiser 2015) defend this claim regarding microbes and microbial communities and Huneman (2014a, b) regarding ecosystems. Other examples are populations, immune systems, and gene regulatory networks. There might also be cases in which the spatial location of, for instance, an organism plays no role in whether or not it is a part of a population (e.g., because the population is said to have a certain spatial distribution). But we still would not say that the population is an object that has a natural boundary inside which its organisms must be located and that can be identified independently of identifying the organisms of the population. To capture cases like these, the condition of spatial inclusion is formulated conditionally.

1.1.2 Temporal Inclusion

The first thing to note is that the relata of temporal inclusion cannot be of the same ontological kind as the relata of spatial inclusion. Spatial inclusion requires its relata to be spatially extended, that is, to be continuants such as material objects,[5] whereas the relata of temporal inclusion must be temporally extended, that is, they must be occurrents such as processes, states, and events. Hence, even if the part-whole relation can be constructed such that it exists between biological *objects*, we cannot understand the conditions under which the part-whole relation holds without also taking into account the biological *processes* in which these objects are involved.[6]

Regarding the whole, I also speak about the behaviors that it characteristically displays. The notion of the characteristic behaviors B_1, … B_n of a biological object Y is, besides the notion of a natural boundary, another central element of my account of biological parthood. I understand the term 'behavior' in a broad way: it refers to what objects do or undergo, to the processes, states, and events in which objects are involved. The underlying idea is that objects in the living world are identified also and in particular by the behaviors they characteristically display. The term 'characteristically' here means that a biological object shows these behaviors under a wider range and variety of contexts in which it is naturally found. The characteristic behaviors of organisms include, for instance, that they live and reproduce, trees grow and convert carbon-dioxide into oxygen (photosynthesize), and cells synthesize proteins, divide into two daughter cells, grow, and differentiate into specific kinds of cells. Which behaviors (or processes) are conceived of as typical or characteristic for a token object depends also on the kind into which this object is classified. Whereas muscle cells, for instance, have the ability to contract if a neuronal signal arrives, blood cells transport oxygen. Frequently, these characteristic behaviors will be instantiations of functional properties (or realizations of functions) of these objects.[7]

It is important to note that most biological objects show *more than one* characteristic behavior (which is a main difference to biological mechanisms; Kaiser and Krickel forthcoming; more on this in Sect. 1.2), and that all of these behaviors B_1, … B_n determine whether an object is a part of the whole Y. For example, it would be inadequate to conclude that centrosomes are no parts of plant cells because they are irrelevant to the biosynthesis of proteins or to photosynthesis. It is also characteristic of a plant cell that it divides into two (if mitosis-promoting factors are present) and the duplication of centrosomes is highly relevant to this process of cell division. The condition of temporal inclusion can thus be specified as follows:

---

[5] At least if one rejects four-dimensionalism (see e.g., Sider 2001).

[6] In what follows, I focus on processes and do not talk about occurrents in general for reasons of simplicity and because processes are of particular importance in the biological realm (Dupré 2012).

[7] However, if one endorses an etiological theory of function (rather than, for instance, a causal-role theory) the class of characteristic behaviors will be larger than the class of realizations of functions (cf. Kitcher 1993). I thus prefer not to speak about functions but rather about behaviors or processes in general.

Temporal Inclusion
Object *X* is a biological part of object *Y* if

(2) *X* is involved in biological process *A*, *Y* characteristically exhibits behaviors B_1,… B_n, and *A takes place during* any of the time periods in which *Y* exhibits B_1, B_2, or… B_n.

Still, there are cases in which the conditions of spatial and temporal inclusion are both satisfied but we would not say that a part-whole relation in fact exists. For example, the cotton ball that a doctor has left inside my abdomen during surgery might soften during the same time period in which I ingest food or move. But this process does not make the cotton ball a proper biological part of me as the softening is irrelevant to me ingesting or me moving. Examples like this suggests that we need to add another criterion to get a sufficient account of biological parthood, namely a relevance criterion.

1.1.3 Relevance

Some analyses of biological part-whole relations and decomposition are not concerned with the notion of a spatiotemporal part. Rather, they focus on *causal interactions* and claim that the different intensities and bandwidths of interaction between objects guide grouping objects together into wholes and decomposing objects into parts. This interactionist approach has been put forward for example by Simon (1962, 1973) and Wimsatt (1976a, 2007).[8] Their central thesis is that the causal interactions among different objects are weaker, fewer, and simpler than the interactions within these objects (i.e., among the parts of each object), which allows different objects to be identified and distinguished on the basis of intensities and bandwidths of interaction (for a recent development of this idea see Huneman 2014a, b).

I think that the interactionist approach draws attention to a significant insight, namely that you have to consider the causal relations among parts (or the causal processes in which parts are involved) in order to group them together as belonging to the same whole *Y* and to characterize them as *parts of Y*. However, in developing my account of biological parthood (Kaiser forthcoming a), I show that the interactionist approach also overlooks two crucial points. First, the biological parts X_1,… X_n of a particular object or system *Y* are not only bound together by the number and intensity of interactions between them. What is even more important is that the parts work together in such a way that the whole *Y* displays its characteristic behaviors B_1,… B_n. This holistic aspect, the reference to the typical behaviors of the whole (i.e., to the causal processes in which the whole is typically involved), is missing in

---

[8] The interactionist approach has been developed further by Haugeland (1998), Grush (2003), and McShea (2000), McShea and Venit (2001).

the interactionist approach.[9] A second aspect that, at least in many cases, binds together biological parts of *Y* is that they are surrounded by a natural boundary such as a membrane (recall Sect. 1.1.1). Natural boundaries spatially constrain causal interactions or processes; that is, they preclude certain interactions (e.g., because the objects are located on different sides of the natural boundary) and allow for or even promote other interactions. The interactionist approach overlooks this spatial constraint.[10]

But how can we spell out the requirement that biological parts work together in such a way that the whole *Y* displays its characteristic behaviors $B_1, \ldots B_n$? How can we specify the claim that the processes in which the candidate parts are involved must be *relevant* to the characteristic behaviors of the whole? In the literature on biological parts and on mechanisms, some suggestions have been made. For instance, Craver (2007a, 141) argues that the components of a mechanism and the behavior of a mechanism must be mutually manipulable. Mellor (2008, 69–71) holds that for *X* to be a part of *Y*, it must have significantly large causal effects on those properties of *Y* that we take to be important to things of that kind. Gillett (2013, 321) claims that for *X* to be a part of *Y*, it must be a member of a team of individuals many of whose members bear powerful and/or productive relations to each other. Jansen and Schulz (2014, 167) have recently pointed out that biologists exclude objects as parts of a whole *Y* for instance on the grounds that these objects do not contribute to or are not necessary to the functioning of *Y*. These different suggestions are critically discussed in Kaiser (forthcoming a). For the purposes of this book it suffices to contrast them with my own proposal for specifying the condition of relevance.

According to my account of biological parthood, a third necessary condition for an object *X* to be a biological part of another object *Y* is that *X* is involved in a biological process *A* that is relevant to at least one of the characteristic behaviors $B_1, \ldots B_n$ of *Y*. This condition can be understood as *X*'s *A* being necessary for at least one of *Y*'s $B_1, \ldots B_n$ (in counterfactual terms: if object *X* had not been involved in process *A*, object *Y* would not have displayed at least one of its characteristic behaviors $B_1, \ldots B_n$). But since this understanding excludes redundant parts such as the second kidney of humans, I prefer to formulate the relevance condition as an INUS condition: the occurrence of process *A* (in which *X* is involved) is an insufficient but necessary member of a collection of biological processes that is itself unnecessary but sufficient for *Y* displaying one of $B_1, \ldots B_n$. This formulation also has the advantage of emphasizing that a particular biological part on its own is insufficient for the functioning or behaving of the whole. In the biological world, wholes are always composed of multiple

---

[9] Note that this holistic aspect does *not* render my account of biological parthood *circular* since one can identify a biological object *Y* and demarcate *Y* from its context by identifying its characteristic behaviors (and by identifying its natural boundary) without, at the same time, knowing what the biological parts of *Y* are.

[10] The conditions of spatial inclusion and of relevance are not completely detached from each other as, for instance, the existence of a natural boundary requires the existence of certain causal cut-offs at that boundary.

objects being involved in multiple processes (i.e., interacting with each other in multiple ways), and this joint working is crucial for the whole displaying its characteristic behaviors. The condition of relevance can be formulated as follows:

Relevance
Object *X* is a biological part of object *Y* if

(3) *A is relevant* to at least one of the behaviors B_1,… B_n that *Y* characteristically exhibits.

One might object that Mackie (1974) has introduced the INUS condition to spell out relations of causal relevance and that using the INUS condition as a condition for part-whole relations confounds the important difference between causal relations and part-whole relations (which are typically assumed to be non-causal). A similar objection has been raised against Craver's mutual manipulability condition (2007a). My response is the following: Even if my condition of relevance turned out to be a condition of causal relevance, this would not imply that part-whole relations are a special kind of causal relations. The condition of relevance is only one of three conditions, and it can be plausibly argued that the other two conditions rule out that part-whole relations are causal relations. For instance, the condition that *X* is spatially included in Y seems to conflict with the assumption that cause and effect must be distinct (and independent) entities. Furthermore, the condition that X's A takes place while Y exhibits B_1, B_2, or… B_n appears to be inconsistent with the assumption that causes precede their effects.

The following overview summarizes my account of biological parthood:

Biological Parthood
Object *X* is a biological part of object *Y* iff

(1) if *Y* has a natural boundary then *X* is *spatially located inside* or *in* the region of the natural boundary (**Spatial Inclusion**),
(2) *X* is involved in biological process *A*, *Y* characteristically exhibits behaviors B_1,… B_n, and *A takes place during* any of the time periods in which *Y* exhibits B_1, B_2, or… B_n (**Temporal Inclusion**), and
(3) *A is relevant* to at least one of the behaviors B_1,… B_n that *Y* characteristically exhibits (**Relevance**).

1.2 Levels as Determined by Part-Whole Relations and Kinds

The idea of reduction is closely connected to the assumption that nature is organized into levels and that reduction relates entities from different levels (or entities that make claims about different levels). This assumption is true for all kinds of reduction (recall Chap. 3, Sect. 3). 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 on a lower level. Theory reduction relates theories about higher-level entities to theories about lower-level entities. An important element of my account of explanatory reduction is that, in the case of a reductive explanation, the phenomenon to be explained is located on a higher level than the explanatorily relevant factors. This element is not all there is for a biological explanation to be reductive, but it is a significant feature.

Levels-talk is not only ubiquitous in philosophy,[11] but also in the sciences themselves. The view that nature is hierarchically organized into a number of discrete levels is deeply embedded in the biological sciences.[12] This is indicated, for instance, by the fact that many biology textbooks start with depicting the living world as being organized into a hierarchy of levels (e.g., Reece et al. 2011, 50f). The depicted levels range typically from molecules and cells over tissues, organs, and organisms up to populations and ecosystems. Despite this ubiquity, the philosophical task of specifying the concept of level has turned out notoriously problematic. Given these problems some authors conclude that the concept of levels should be abandoned altogether (Guttman 1976, Potochnik and McGill 2012, Eronen 2015). Others adhere to it, but shy away from specifying the notion of level such that it refers to global levels of organization. Instead, they suggest conceiving levels as being *local* levels that are defined only relative to particular explanatory interests (Craver 2007a, Love 2012b). One of the most elaborated conceptions of levels, the mechanistic account, argues in favor of the locality of levels. Since, in the philosophy of the life sciences, the mechanistic account is prevalent and since it also is an important point of reference of my own account of levels, I will briefly introduce the concept of mechanistic levels before I sketch the main theses of my own approach.

The mechanistic account of levels has been put forward primarily by Craver (2007a, Chapter 5) but can also be found in the work of Bechtel (2008, 141–148; Bechtel and Hamilton 2007, 411–413). A central claim of Craver's account is that levels of nature are determined by a special kind of part-whole relations, namely by the relations between the *components* of a mechanism (which are said to be "acting entities" 2007a, 189) and the (behaving) *mechanism as a whole*. According to the mechanistic approach, levels of mechanisms are only defined *locally*, that is, only within a given compositional hierarchy (between a certain mechanism and its components); there exist no monolithic levels of nature as the ones that, for instance, Wimsatt envisions (1976b, 2007).[13] It does not make sense, so the mechanists, for

---

[11] In reference to Simon's classic 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). For a rejection of this view see Ladyman and Ross 2007, 55–57.

[12] Hence, I understand levels of organization to be *levels of nature*, not levels of science (as e.g., Oppenheim and Putnam 1958 did). That is, levels relate *entities in the world* such as objects, properties, or processes, not epistemic entities such as theories, explanations, or scientific disciplines.

[13] Wimsatt characterizes a level of organization as a "local maximum of predictability and regularity" (1976b, 238).

example, to ask 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 the liver's behaviors. 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" (Craver 2007a, 192).

The account of levels of organization that I develop (Kaiser manuscript a) deviates from the mechanistic approach in several important respects.[14] A major objection to the mechanistic account is the following: Since it traces back the hierarchical organization of the living world to component-mechanism relations, it results in a view of levels that is far *too local* and *too restricted*. It thus fails to capture how levels are understood in many contexts in the biological sciences. I adopt a twofold strategy to elude this objection and to develop an account of levels that is less restricted and more global. First, I claim that levels of organization are not determined by component-mechanism relations but rather by *part-whole relations* in general (which should be conceived of in accordance with my account of biological parthood; see Sect. 1.1, Kaiser forthcoming a). Parts are located on a lower level than the wholes they are parts of. The decomposition of a biological object or system *Y* into its parts $X_1, \ldots X_n$ thus comprises identifying at least two levels of organization: the level on which the parts are located and the higher level on which the whole is located.[15] Second, I point out that part-whole relations are not all there is to distinguishing levels of organization. There is a second major factor that determines the level on which a biological object is located, namely the *membership* of an object to a certain *general biological kind*. Accordingly, I identify two sufficient conditions under which one object is located on a lower level than another object:

Levels of Organization

Object *X* is located on a lower level than object *Y* iff

(1) *X* is a *biological part* of *Y* **(Biological Parthood)** *or*

(2) *X* belongs to the same *general biological kind* as one or more of the biological parts of *Y* **(Kind Membership)**.

These two conditions reveal in which way my view of levels differs from and is superior to the mechanistic account. First, there are two main differences between a part-whole relation as I understand it and the relation between the components of a

---

[14] I agree with Craver on the point that levels always are, as he frames it, "levels of behaving components" (2007a, 190) and not only levels of objects. In my terms, since levels are determined by part-whole relations, which cannot be identified by considering only objects (processes must be taken into account as well), not only are objects located on levels but objects that are involved in certain processes or that behave in a certain way.

[15] There might be additional reasons not to locate all of Y's parts on the same level of organization (see the second point).

mechanism and the mechanism as a whole (cf. Kaiser and Krickel forthcoming). The first difference concerns the generally accepted claim that mechanisms always are mechanisms for *one* specific *behavior* (e.g., Glennan 2002, 344; Craver 2007a, 122–128). Mechanisms are said to be composed only of those entities (i.e., objects) and activities (i.e., occurrents or processes) that are relevant to this one behavior. Take the example of the mechanism by which a paramecium (i.e., a unicellular ciliate protozoon) gathers food. This mechanism consists, for instance, of cilia that stroke and hereby sweep up food into the cell mouth of the paramecium. According to the mechanistic account, claims about levels can only be made with respect to a certain mechanism and the components that are relevant to the one behavior related to this mechanism. The cilia and the cell mouth can be said to be located on a lower level than the paramecium. But other parts of the paramecium, such as its micronucleus, its contractile vacuole, or its cell membrane, cannot be said to be located on the same level as the cilia and the cell mouth because they are not components of the mechanism for food gathering.

Hence, the mechanistic account of levels allows only for a very limited range of level claims. This is an implausible consequence since biologists are far more permissive in the claims they make about levels. They do not restrict their level claims to mechanisms and their components. Since the biological parts of an object *Y* are individuated not only with respect to one behavior of *Y* (for which there might be a mechanism) but with respect to all of the *many* behaviors that *Y* characteristically exhibits an account that regards levels as being determined by part-whole relations (instead of component-mechanism relations) avoids this implausible restriction and allows for claims such as the following: The moving cilia (which are relevant to food gathering), the duplicating micronucleus (which is relevant to asexual reproduction), and the contracting vacuole (which is relevant to osmosis) are all located on a lower level of organization than the paramecium.

A second main difference between part-whole and component-mechanism relations is that natural boundaries of biological objects (wholes) are crucial to identifying the parts of these objects. By contrast, the mechanisms' boundaries frequently do not correspond to existing natural boundaries. In accordance, Craver states that "mechanisms frequently transgress compartmental boundaries" (2007a, 141). In the mechanistic view, mechanisms are demarcated from their context simply by identifying their components, that is, by identifying all those acting entities that are "constitutively relevant" (Craver 2007a, 139) to the behavior of the mechanism. This view implies that mechanisms often include object- or system-external objects and processes, which calls into question Craver's central thesis that constitutive relevance can be explicated as mutual manipulability plus parthood (cf. Kaiser forthcoming a). This problem at least indirectly challenges the mechanistic account of levels.

Second, biological practice often works with the notion of levels of organization not being restricted to the parts of an object *Y* being located on a lower level than *Y* itself. In particular in biological contexts in which questions about reduction and reductionism are disputed, levels are often conceived of as not being restricted to local explanatory contexts but as dividing the natural world more globally.[16] For example, biologists state that the molecules involved in the intracellular signal transduction process (e.g., G proteins or kinases) are located on the same level as an *extracellular* signal molecule (e.g., an insulin-like growth factor, IGF) that binds to a cell surface receptor and induces an intracellular signal transduction pathway that in the end stimulates for instance cell survival and growth. IGF is located outside the cell membrane and therefore is not a biological part of the cell, whereas G proteins and kinases are biological parts of the cell (they are located inside the cell membrane and they are relevant to signal transduction). However, biologists do not only claim that G proteins and kinases (parts of the cell) are located on a lower level than the cell, they also claim that IGF (not a part of the cell) is located on the same lower level than the cell.

In my view, a philosophical notion of level must account for level claims of this kind, which are widespread in the biological sciences. What is more, the account of reductive explanation that I develop requires a view of levels that characterizes levels more globally. For example, the prevalent idea that there is a single level on which all molecules are located (and which is regarded as the most fundamental biological level) does not make sense if levels are supposed to be defined only locally (more on this in Sect. 2.4.1). Furthermore, one of my central claims will be that, for the reductive character of an explanation, it matters on which level environmental factors that are not parts of the object in question are located. This is possible only if part-whole relations are not the only factor that determines levels. The additional, second factor that I identify is the membership of an object to a certain *general biological kind*, which allows non-parts to be included into level claims and different particular part-whole relations to be integrated into hierarchies of part-whole relations (cf. Fazekas and Kertész 2011, 379f). The previous example illustrates this: IGF, G proteins, and kinases are said to be located on the same level because all of them are macromolecules. IGF is located on a lower level than the cell not because it is a part of the cell but because it belongs to the same biological kind as other parts of the cell (e.g., G proteins and kinases).

Let me conclude this overview by pointing out an important consequence of my account of what levels of organization are and how they are determined. My account implies that the set of all factors that are located on a lower level than the object Y in question is *larger* than the set of factors that are internal to Y, that is, that are biological parts of Y. This implication will turn out to be important in distinguishing the first and second feature of reductive explanations, namely their lower-level character (Sect. 2) and their internal character (Sect. 3).

---

[16] This does not mean that these biologists accept the simple and, admittedly, unconvincing view that everything in nature is organized into a few, monolithic levels. An intermediate position is possible that rejects the radical view that levels are defined only in local explanatory contexts and conceives levels to be more global (but not too global) things.

1.3 The Methodology of My Account

As I pointed out in Chap. 2, the goal of this book is to provide an understanding of what reductive explanation in contemporary biological practice is. Accordingly, I accept descriptive adequacy as an important criterion of adequacy and develop my account in a bottom-up fashion, which means that it emerges from a critical reconstruction of the reductive reasoning and investigative practices that are present in contemporary biology. The critical-normative character of my analysis is apparent in three features. First, my analysis does not consider all available cases of reductive explanation, but rather focuses on paradigmatic and important examples of reductive explanations. Second, it explicates assumptions about reductive explanations that are only implicitly present in biological practice. Third, in order to construct a coherent and (to a certain degree) universal account, I need to assess some biologists' statements as incorrect, too vague, or inappropriate, and exclude them from the empirical basis of my account. The goal of this critical reconstruction is twofold. 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. Furthermore, I critically reconstruct 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.

Finally, let me elaborate on the question of what these empirical data are, on which my analysis is built (see also Chap. 2, Sect. 1.4). In my view, taking biology seriously implies not merely discussing old standard philosophical examples, such as Mendel's second law of independent assortment or the cubical-peg-square-hole example, again and again. Instead, philosophers should analyze the various reductive reasoning practices that are present in actual biological research and the way they are conceptualized in biological discussions about reductionism (if they are). 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.[17] These examples are, in part, those that are discussed in the philosophy of biology. However, the majority of the examples are taken from the biological literature itself. I assume that paradigmatic examples are those that are representative of a certain field, that are long-standing and much discussed, and that thus may have found their way into a biology textbook. Important examples are those that are, for instance, regarded as successful science or that contribute to achieving a certain aim of the field (e.g., manipulation or medical treatments). Examples might also be important because they are particularly relevant to a certain philosophical question, such as molecular

---

[17] Since I want my account to apply to biology in general and not to be restricted to certain biological disciplines, the strategy Sarkar (1998), and 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.

biological explanations are important for an analysis of reduction because molecular biology has been regarded as the embodiment of the success of reductionism.[18] 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 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 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

---

[18] 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 influence does not result in imposing an ill-fitting view of reduction on biological practice.

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 Chap. 3, Sect. 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 “reductionism 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.

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 biological object or system *Y* by *appealing only to lower-level factors*, that is, to factors that are located on at least one level below the level on which *Y* is located. 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 a biological discipline that is regarded as the embodiment of the success of reductionism, namely molecular biology (Sect. 2.1). In Sect. 2.2 I point out that reductive explanations 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 2.3 serves to specify what exactly it means that biological explanations refer only to lower-level factors. What are these higher-level factors that are ignored in reductive explanations in biology? I briefly discuss whether the distinction between lower-level and higher-level factors coincides with the distinction between parts and non-parts (Sect. 2.3.1). Then I turn to functional properties which might be seen as an obstacle for reduction (Sect. 2.3.2). A popular argument against reductionism states that a whole is more than the sum of its parts as the way the parts are organized matters to the identity of the whole. This raises the question of whether organizational features of parts really count as higher-level factors. In other words, does reference to the way parts are spatially and temporally organized into a whole violate the lower-level character of an explanation (Sect. 2.3.3)? In Sect. 2.4 I introduce two significant subtypes of lower-level explanations, namely fundamental-level explanations, which I characterize as molecular explanations that encompass genetic explanations (Sects. 2.4.1 and 2.4.2), and single-factor explanations (Sect. 2.4.4). 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 (Sect. 2.4.4).

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.[19] 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 and Tabery 2010), but it has also been expanded to many other biological fields. Cell biology, developmental biology, genetics, and immunology are just some of the disciplines that "went molecular" (Darden and Tabery 2010). Nowadays molecular techniques are utilized in almost all biological disciplines.

In the 1970s and 1980s 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

---

[19] Also in biology itself molecular biology is frequently characterized as the "triumph of reductionism" (Fang and Casadevall 2011, 1402).

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 Sect. 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 Chap. 4, Sect. 1.3 and in Chap. 5, Sect. 4 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 Fig. 6.1 indicates.

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 thyla-

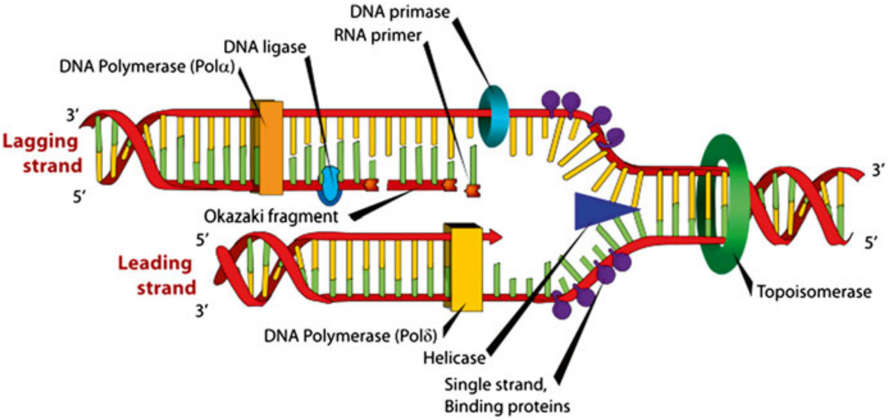

Fig. 6.1 DNA replication

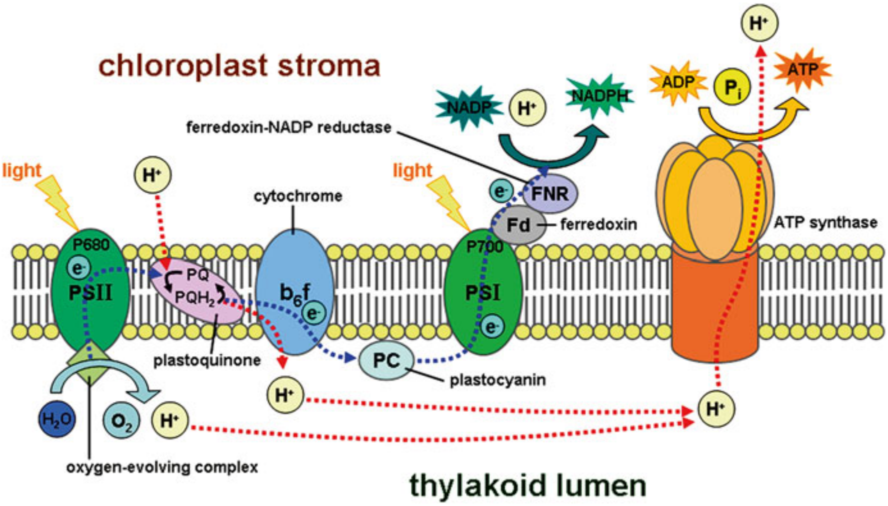

Fig. 6.2 Light reaction of photosynthesis

koid membrane in the chloroplasts (see Fig. 6.2). 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

What are the features due to which these two explanations from molecular biology are characterized as reductive? One answer is that these explanations both explain a particular phenomenon (i.e., a behavior of a biological object or system) 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 than the former (the concept of a level is understood according to my account of levels of organization (Kaiser manuscript a, recall Sect. 1.2). For instance, the behavior of a cell to duplicate its genetic information is reductively explained by describing the organization of and interactions between certain parts 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 factors, such as photosystems, electrons, electrochemical proton gradients, and ATP synthases. Reductive explanations thus *bridge* at least *two levels*, namely 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 (this is why reductive explanations are also called interlevel explanations; cf. Brigandt and Love 2008, Section 5; Bechtel 2008, 148).[20] 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.[21] One may argue, for instance, that a photosystem is located on a “higher” lower level than an electron because photosystems are, among other things, composed of electrons and because photosystems belong to the kind “macromolecular complex”, whereas electrons are subatomic particles (recall Sect. 1.2). However, unlike the lower-level character, the explanans of reductive explanations *need not* involve different lower levels.

2.2 Unidirectional Flow of Explanation

Another way to express the lower-level character of reductive explanations is to say that the direction of reductive explanations runs from the lower level to the higher level, not also in the opposite direction. This is why biologists also refer to reductive

---

[20] 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 (1958). My account of reduction is neither concerned with theory reduction (recall Chap. 3, Sect. 3.1), nor does it accept Oppenheim’s and Putnam’s concept of level.

[21] This is commonly referred to as the “multilevel character” of explanation (e.g. Mitchell 2009, 109–115; Schaffner 2006, 384–387; Craver 2007a, 9–16).

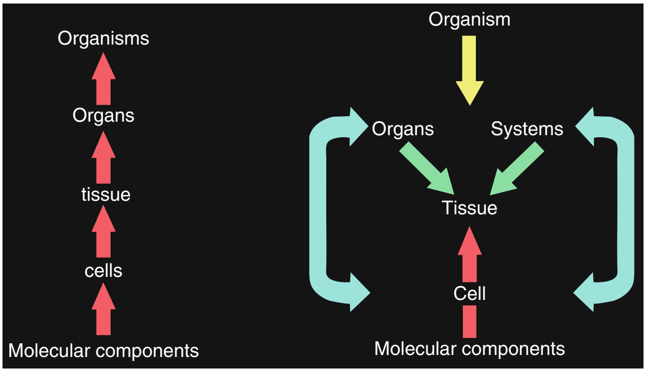

Fig. 6.3 Reductionist vs. systemic approaches in cancer research (Reprinted from Soto et al. 2009, 6; with permission from Elsevier)

explanations as "*bottom-up [explanations]*" (Sorger 2005, 9; Soto et al. 2009, 5; my emphasis) or "upward explanation[s]" (Mikkelson 2004, 120; Lidicker 1988, 278).[22] The 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 reveals 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 level) to the bottom (i.e., lower level).[23] In cancer research this difference between, as they call it, the "reductionist" and the "organicist/systemic approach" (Soto et al. 2009, 6) is illustrated by Fig. 6.3.

According to Soto et al., reductionists assume that causal relations between entities from different levels exclusively run from lower to higher levels, whereas proponents

---

[22] Note that the direction of a reductive explanation runs in the opposite direction than the reductive method decomposition. Biological objects or systems are decomposed into their parts by looking down one or more levels, whereas the behavior of an object is explained by looking upward.

[23] 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)

of an organicist/systemic approach allow causal relations also to run from higher to lower levels (or in both directions).[24] 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 and Bechtel 2007). It is thus questionable whether the arrows in the above figure can be interpreted as causal relations. Instead, one could argue that they represent part-whole relations (the red arrows) or relations 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. But 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 part-whole relations suffices, too).

Despite the problematic reliance on downward causation, Soto's et al. argument also has a virtue. It discloses a *possible line of criticism* against reductive explanation. A common antireductionist objection is to argue that adequate explanations of (many or certain) biological phenomena do not exhibit a unidirectional flow of explanation. That is, antireductionists frequently claim that adequate explanations of biological phenomena also flow from higher to lower levels, not only 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).

2.3 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. But what exactly does it mean that higher-level factors are excluded from the explanans? Can descriptions of higher-level factors simply be replaced by descriptions of lower-level factors (Sect. 2.3.2)? How can the distinction between lower-level and higher-level factors be drawn? Which kinds of factors belong for instance to the set of lower-level factors: only parts (Sect. 2.3.1), also functional properties of objects (Sect. 2.3.3), and also the organization of and interaction between parts (Sect. 2.3.4)?

---

[24] 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)

2.3.1 Why Lower-Level Factors Are More Than Parts

This point will be discussed in detail in Sect. 3.2, which is why I will only briefly examine it here. One might be faintly reminiscent of my account of levels (recall Sect. 1.2; Kaiser manuscript a) and propose to identify a higher-level factor with the biological object (the whole) whose behavior is to be explained and lower-level factors with the biological parts of this object. It is true that part-whole relations are one major factor that determines levels. However, as I have pointed out in Sect. 1.2, there is a second factor that also determines on which level a certain object is located (relative to another object *Y*), namely whether *X* is a member of the *same general biological kind* as at least one of *Y*'s biological parts.

An important consequence of my claim that kind membership co-determines level distinctions is that not only biological parts can be said to be located on a lower level of organization than their whole *Y*. Also objects that are not parts of *Y*, but that belong to the same general biological kind as at least one of *Y*'s parts can be said to be located on a lower level than *Y*. Hence, the set of all lower-level factors (where 'lower' is understood relative to the level of Y) is *larger* than the set of all biological parts (of *Y*). The difference between lower-level and higher-level factors thus only partially coincides with the distinction between parts and non-parts.

2.3.2 Are Higher-Level Terms Dispensable?

Let us assume for the sake of argument that the term 'being located on a lower level than object *Y*' has the same extension as the term 'being a biological part of object *Y*'. Even if we accepted this, the question of what these lower-level factors are that are exclusively referred to in reductive explanations would still not be fully answered. For instance, it would remain an open question whether all kinds of properties of biological 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.[25] 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

---

[25] 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)

> 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). 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.[26]
>
> 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, that is, 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, the 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').[27] In particular,

---

[26] For more information about limb development see Gilbert 2006, Chapter 16.

[27] 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 Chap. 4, Sect. 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 Chap. 4, Sect. 1.1.2).

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 as Kitcher might counter that an explanation that 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 Chap. 5, Sect. 4). My central concern is rather how these authors spell out which kinds of objects and properties are lower-level factors and, thus, allowed in (the explanans of) reductive explanations and which of them count as higher-level factors. 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 *(macro-) molecular level*. Accordingly, if the explanans 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). But Rosenberg's and Kitcher's discussion does not help to answer the question of whether the organization and functional properties of parts also count as lower-level. This question is addressed in the next two sections by considering the contributions that other authors have made to this discussion.

2.3.3 Do Functional Properties Count as Lower-Level?

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 and Wagner 2001; Frost-Arnold 2004) tackle 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.[28] Accordingly, the question is whether we can refer to functional properties of 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.[29] 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 and 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.[30]

---

[28] 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 biological parts of the blastula and located on a lower level than it, but they are not molecules.

[29] 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." (2004a, 1016; see also van Regenmortel 1998)

[30] 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.

But 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 Sect. 3 when I analyze the internal character of reductive explanations. Here the *context of the parts* of the object *Y* to be explained is meant, which may include only factors that are internal to Y, that is, that are also parts of *Y*, or which may (but need not) also include parts of the *context of the object Y* (the whole) itself. If the functional properties of *Y*'s parts are specified by appealing to lower-level factors that are external to (i.e., not part of) *Y*, the lower-level character of this explanation will not violated, but its internal character will be (for further details see Sect. 3). Second, it is important not to confuse two questions: On the one hand, there is the question that is at issue here, namely the question of whether the inclusion of functional properties of the parts of a biological object *Y* violates the reductive character of the explanation of the object's behavior. This question should be kept apart from another question, namely the question of whether functional properties of biological objects itself can be explained in a reductive manner or not.[31] An affirmative answer to the former does not imply any kind of answer to the latter. For example, it may be that we can reductively explain the behavior of the human heart to contract and that this is compatible with describing also functional properties of the biological parts of the heart (e.g. the function of the Purkinje fibers to transmit the electric charge to the myocardium). But whether the function of the heart to pump blood can be explained reductively or whether it must prominently refer to environmental factors of the heart (e.g. the organism as a whole) is a different issue.

2.3.4 Is the Organization of Parts a Lower-Level Factor?

Finally, let us discuss whether the *organization* of the biological parts of an object *Y* and the *interactions* between the parts of *Y* also count as lower-level factors. 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"[32]) of parts into an object as a whole is a common line of criticism of antireductionists. But 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

---

[31] With regard to the latter question see, for instance, Sarkar 2005, 117–143.

[32] "[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)

approached in detail in Sect. 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. 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 *Y*'s biological parts) or as properties of the *biological object Y* whose behavior is to be explained (i.e., as properties of the whole). Lower-level properties can be referred to in a reductive explanation, whereas reference to properties of the whole *Y* 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 located at the bottom of the embryo'). Only the latter of them are properties of the whole *Y* whose inclusion renders an explanation non-reductive.

To summarize, a central aim of this section was to state more precisely what lower-level factors are, in particular whether functional properties of biological parts and the organization of them belong to the class of lower-level factors. The result of my analysis is threefold: first, the set of factors that are located on a lower level of organization than a particular object *Y* of interest is not restricted to biological parts of *Y*. Rather, factors that are external to *Y*, but are located on the same level(s) than at least one of *Y*'s parts, are also lower-level. 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 biological parts of an object *Y* can be conceptualized as lower-level (relational) properties and not as properties of Y (the whole), it will also count as lower-level factor.

2.4 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. Another important result of my analysis, however, will be that reductive explanations *need not* go down to the lowest biological level of molecules[33] (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 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).[34] 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.[35] Brandon has called the reductionist

---

[33] 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). But 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.

[34] 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 and 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 and Lewontin 1980, 51)

[35] The version of reductionism that corresponds to reductive explanations of this kind is also known as "*methodological individualism*" (Sarkar 2009, Sect. 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.

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 Chap. 5, Sect. 4, 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 also gives rise to a *one-dimensional view of reductive explanation*. 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 Chap. 5, Sect. 4). To put it into the words of some biologists: "in practice, the reductive thrust stops at the level where 'it makes sense'" (Soto and Sonnenschein 2006, 362).[36] In what follows I first explain the notion of a fundamental-level explanation in more detail (Sect. 2.4.1) and discuss genetic explanations as an important example of fundamental-level explanations (Sect. 2.4.2). Then I argue why reductive explanations should not be identified with fundamental level explanations (Sect. 2.4.3). Finally, I identify another subtype of reductive explanation in the biological science, namely single-factor explanations (Sect. 2.4.4).

2.4.1 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'[37]). 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 and 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 and Sonnenschein 2010, 364; my emphasis)

This understanding is especially common when biologists criticize reductionist approaches:

---

[36] 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., 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).

[37] 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).

> 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 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). I call reductive explanations that appeal only to factors that are located on this single fundamental level of molecules (i.e., molecular explanations) *fundamental-level explanations*.[38]

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 (recall Chap. 4, Sects. 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, that is, 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.

2.4.2 Genetic Explanation

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 and Sonnenschein 2005, 104; Soto and Sonnenschein 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 1980s 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. But 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:

---

[38] 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).

> 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 and 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).[39] 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. What is more, during the last decade it has generated one of the most 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.

Ana M. Soto and Carlos Sonnenschein are among the central figures in this dispute (e.g., Sonnenschein and Soto 2000, Soto and Sonnenschein 2005, 2006, 2010, and Soto et al. 2009). 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 and 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 and Sonnenschein 2005, 103), rather than a cellular problem caused by mutated genes.[40]

Let us go into some more 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 and 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

---

[39] 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).

[40] 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" (Sonnenschein and Soto 2000).

> 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.[41]
>
> This is the view of carcinogenesis that is still prevailing in contemporary biomedical research. But 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.[42] The following quotations express this:
>
> 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 and Sonnenschein 2005, 112). This is why they claim that the genesis of cancer is only adequately explained in a non-reductive fashion in terms of the disruption of tissue organization.[43]

The example of cancer research is 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

---

[41] 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).

[42] 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 and Weinberg 2002; Malaterre 2007, 59). However, they do not regard changes at the tissue level as being important, as proponents of TOFT claim.

[43] 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).

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 (recall Chap. 4, Sect. 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 (cf. Waters 1994). Accordingly, genes can be conceived as a subtype of molecules and genetic explanations as a *subtype of molecular explanations*. Genetic explanations thus are cases of fundamental-level explanations.[44]

2.4.3 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[45] results from relative sliding between so called thick filaments (consisting mainly of myosin) and thin filaments (consisting mainly of actin) (Huxley and Niedergerke 1954; Huxley and 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.

According to the *sliding-filament model*, which is depicted in Fig. 6.4, 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.

---

[44] Genetic explanations can also be characterized as single-factor explanations (see Sect. 2.4.3). This is particularly true for genetic explanations that appeal to single genes or single kinds of genes.

[45] 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.

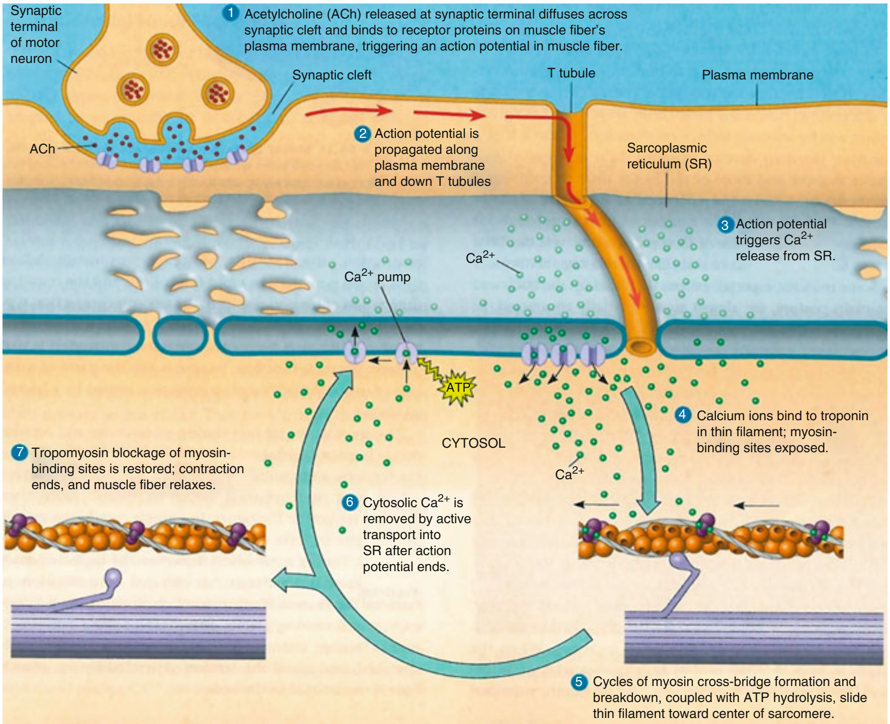

Fig. 6.4 Mechanism of muscle contraction (Reprinted from Reece et al. 2011, 1153; with permission from Pearson)

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.

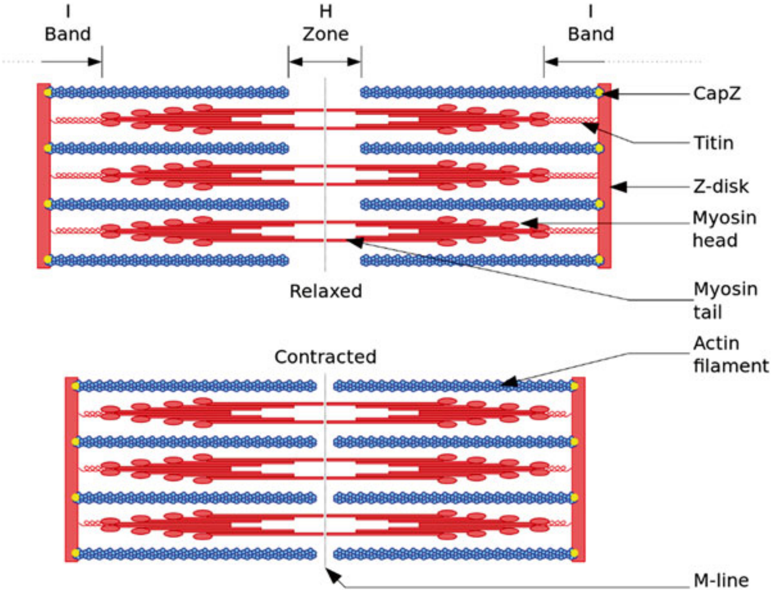

Fig. 6.5 The sarcomere of a skeletal muscle (Richfield 2014)

> 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.
>
> 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, almost disappear (see Fig. 6.5). 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 biological parts of muscle fibers 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 also be applied to the latter, non-molecular explanation.

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. But 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.

2.4.4 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 2004a, 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.[46]

It is important to note that single-factor explanations rarely occur in biological research practice.[47] 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 2004b; 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 fundamental-level explanations; recall Sect. 2.4.2). 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 Focusing on 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 biological object or system *Y* by focusing on those objects (and processes or interactions) that are

---

[46] 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 character of reductive explanations (Sect. 3). I admit that one could allocate this subtype differently. But I think that it suits this section well because the single factor is chosen among the lower-level factors.

[47] 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."

internal to *Y*, that is, that are biological parts of *Y*. In other words, I claim 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. The internal character of a biological explanation is thus just a *typical feature*, but not a necessary condition for the reductivity of an explanation.

I proceed as follows. In Sect. 3.1 I elaborate on the internal character of reductive explanations, first, by analyzing how this feature is conceptualized in biological research papers that discuss reductionism (Sect. 3.1.1), and second, by examining an instructive example of how an explanation fails to exhibit an internal character (Sect. 3.1.2). I then point out the difference between the internal character and the lower-level character of an explanation (Sect. 3.2). In Sect. 3.3 I specify what it means that an explanation “focuses on” factors that are internal to the system of interest. I argue that this amounts to either ignoring environmental factors altogether or simplifying them in a specific manner.

3.1 The Internal Character of Reductive Explanations

3.1.1 “Looking Inward” vs. “Putting Things in a Context”

In the biological literature a second feature of reductive explanations is mentioned – far from as frequently as the lower-level character, but yet frequently enough to be important to my analysis. This feature is implicit in how biologists specify the reductionist methodology. When biologists discuss for instance 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 and 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 object or system *Y* with a reductionist methodology they focus on factors that are *internal* to *Y* and ignore or simplify factors that belong to the environment or context of *Y* (i.e., *external* factors). In contrast, a non-reductionistic proceeding also takes into account the context of the object or system *Y* and, for instance, explores how changes in the environmental conditions affect changes in the interactions between *Y*’s parts and how this in turn influences whether *Y* displays the behavior in question or not. This reductionist research strategy can also be characterized as shielding an object or 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.[48] But since biological objects or systems generally are open systems as well as objects that evolved under the influence of specific environmental conditions, in most cases the environment is crucial for the functioning of these objects or systems and cannot be as easily ignored and simplified as the reductionist might wish.[49] The more important the environment is the more is the adequacy of the reductionistic methodology called into question.

From this I infer that reductive explanations in the biological science exhibit a second characteristic: they explain the behavior of a biological object or system *Y* by focusing on those objects and processes that are *internal* to *Y* (this is called the internal character of a reductive explanation). As I have already pointed out in Sect. 1.1, I understand the term 'internal' in the following sense: a factor *X* is internal to an object or system *Y* (the whole) if *X* is a *biological part* of *Y*. The close connection between the two relations 'being internal to' and 'being a biological part of' is explained by the fact that biological parthood requires spatial inclusion: an object *X* is a biological part of another object *Y* (the whole) if *X* is spatially included in *Y*. The condition of spatial inclusion can be specified as follows: if *Y* (the whole) has a natural boundary then X must be spatially located inside or in the region that *Y*'s natural boundary occupies (Kaiser forthcoming a; recall Sect. 1.1). Even though spatial inclusion is not the only requirement for biological parthood (the conditions of temporal inclusion and of relevance must be satisfied as well) it explains why 'being a biological part of' is so closely connected to 'being internal to'. Because everything that is internal to an object *Y* is at the same time a biological part of *Y* (and the other way round), the set of all explanations with an internal character is identical to the set of all *part-whole explanations* (remember that this is not true for lower-level explanations and thus for reductive explanations in general; more on this in Sect. 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 (i.e., the behavior of an object) reductively implies to *look inside* the object or system in question, not outside of it (likewise, it implies to look at lower levels of organization than the level of the whole, 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 has not been recognized in the philosophical literature for a long time. If the internal character were mentioned at all, it has been lumped

---

[48] Bechtel and Richardson put a lot of effort in depicting these "psychological constraints" for developing explanations (see e.g., 2010, 234–243).

[49] 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).

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.[50] In addition, Hüttemann and Love (2011) point out that the internal character and the lower-level character – they call them "intrinsicality"[51] and "fundamentality" (2011, 527) – are two distinct "aspects" (2011, 523) of reductive explanations (recall Chap. 4, Sect. 2.3).

3.1.2 When the Focus on Internal Factors Fails

Just as an explanation can fail to possess a lower-level character it can also fail to exhibit an internal character. This is the case when certain contextual factors are so crucial to the behavior of the considered biological object 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).

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[52] 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

---

[50] 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)

[51] The term 'intrinsic' might have the advantage (compared to 'internal') that it can also be spelled out in a non-spatial way and thus captures also more rare types of reductive explanations (e.g. structural explanations). However, I stick to the term 'internal' because my view of part-whole relations is closely linked to the idea that parts are spatially included in their wholes. Moreover, the term 'intrinsic' has the drawback that it confounds the internal character with the third characteristic of reductive explanations, which is introduced in Sect. 4.

[52] 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).

> 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.[53] Such an explanation would be reductive in two respects: first, it would be reductive because it would exclusively refer to interacting objects (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 objects (i.e., amino acids) that are internal to the object whose behavior is to be explained (i.e., the polypeptide or protein).

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 (see Fig. 6.6). 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 and Hayer-Hartl 2002; Dobson 2003; Frydman 2001).

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). The

---

[53] This thesis is sometimes referred to as (one reading of) the “linear sequence hypothesis” (Hüttemann and Love 2011, 17), in short “LSH”.

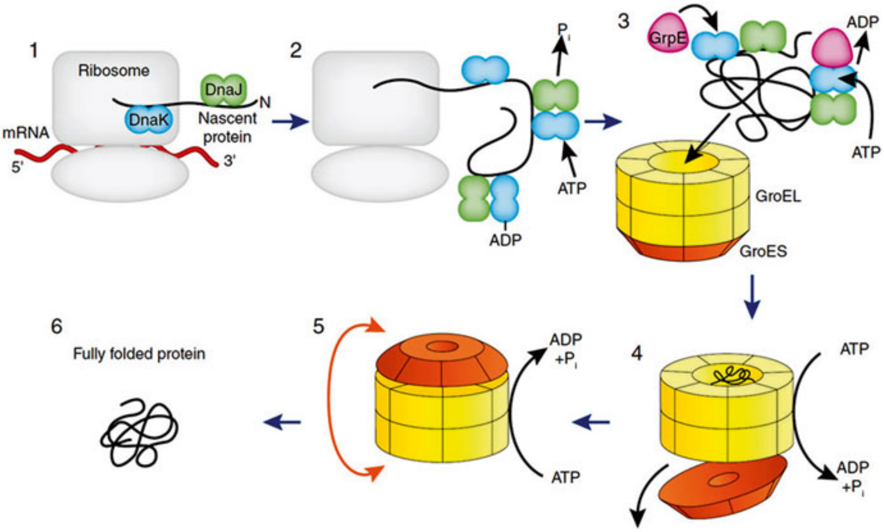

Fig. 6.6 Chaperone-assisted protein folding (Reprinted from Hartl 2011, 1208; with permission from Nature Publishing Group)

reductionist strategy of excluding chaperones from the explanation of protein folding or of simplifying them as mere background or input conditions (more on this in Sect. 3.3) thus 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 object *Y* under study reaches its limits. In cases like this the environment of *Y* 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. This is due to my account of levels of organization and, in particular, to my thesis that part-whole relations are one central factor that determines levels (Kaiser manuscript a; recall Sect. 1.2). Any factor that is internal to a particular object *Y* (the whole) is a biological part of *Y* and, thus, automatically located on a lower level of organization than *Y*. Accordingly, if an explanation exhibits an internal character, it also displays a lower-level character. To put it another way, the set of all explanations with an internal character (i.e., the set of all part-whole explanations) is a *subset* of the set of all lower-level explanations.

The term 'subset', however, implies that there exist lower-level explanations that do not possess an internal character. That is, there exist reductive or lower-level explanations which are not part-whole explanations. Hence, 'exhibiting an internal character' must be distinguished from 'exhibiting a lower-level character'. This is a crucial difference between my account and Hüttemann's and Love's analysis (2011): they identify reductive explanation with part-whole explanation, whereas I distinguish these two kinds of explanation (recall Chap. 4, Sect. 2.3). As I have already mentioned, it is possible that explanations appeal to environmental factors (i.e., non-parts) and, nevertheless, are lower-level explanations since they refer only to those objects and processes in the environment of an object *Y* that are located on a lower level than *Y*. This is possible because a part-whole relation between two objects *X* and *Y* is not the only factor that determines whether *X* is located on a lower level than *Y*. Even if *X* were not a part of *Y* it would be located on a lower level of *Y* if it belonged to at least one of the general biological kinds to which the biological parts of *Y* belong. Hence, an object need not be internal to *Y* (i.e., be a part of *Y*) to be located on a lower level of *Y*.

Recall, for instance, the explanation of muscle contraction (Sect. 2.4.3). This explanation refers to, amongst other things, an incoming neuronal signal which initiates the contraction process. This "incoming neuronal signal" means that signal molecules (in this case the neurotransmitter acetylcholine) are released into the synaptic cleft and bind to receptors at the outer side of the muscle fiber membrane. These neurotransmitter molecules belongs to the environment or context of the muscle fiber and are not parts of the muscle fiber as they are located outside of the natural boundary of the muscle fiber, namely it's cell membrane. The explanation of muscle contraction thus does not possess an internal character since it prominently refers to external factors (the neurotransmitters). Yet, the explanation is a lower-level explanation since it refers only to factors that either are biological parts of the muscle fiber (e.g. calcium ions, actin filaments, tropomyosin molecules, etc.) or that are external to the muscle fiber but belong to the same general biological kind as one of the parts of the muscle fiber do, namely to the kind of molecules (e.g. neurotransmitter molecules).[54]

---

[54] Another example of a lower-level explanation that does not also exhibit an internal character is the explanation of how the PI 3-kinase/Akt pathway promotes cell survival (see Sect. 3.3). This explanation also describes how an extracellular survival signal binds 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 kind of macromolecules, to which also most of the cell's parts belong.

Reductive explanations of this kind – 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 also their context. A reductionist may counter that reductionism can be preserved by reducing also the context. The phrase 'reducing the context' means that the context 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 explanations with an internal character (i.e., of all part-whole explanations) 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 on' 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.[55] If we, however, analyze paradigmatic cases of reductive

---

[55] I only found two statements of biologists that are relevant to these questions. But even 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)

explanations 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 types 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),[56] 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)

"To make a multicellular organism, cells must communicate, just as humans must communicate if they are to organize themselves into a complex society." (Alberts et al. 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 in Fig. 6.7. 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)P3, 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,

---

[56] Machamer et al. speak about start or set-up conditions, too (2000, 11). There is, however, an important difference between our positions: they regard these conditions to be components 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 biological object *Y*, not as biological parts of *Y* itself.

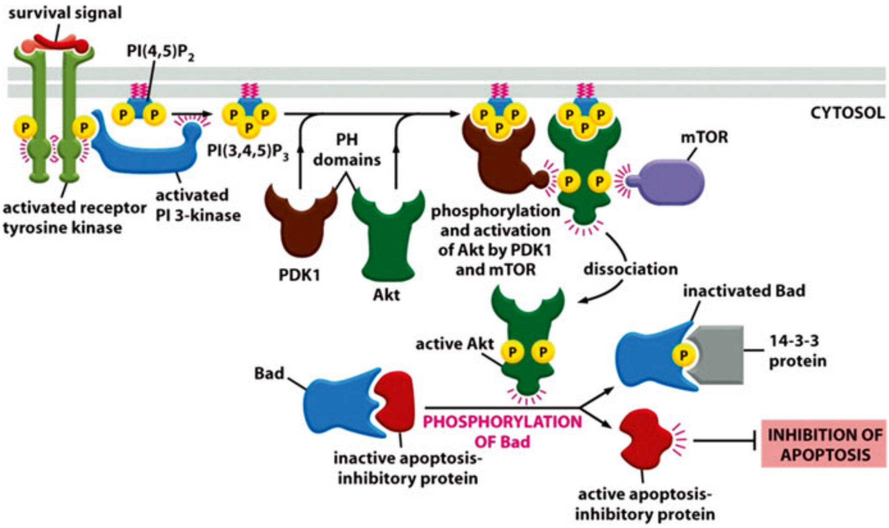

Fig. 6.7 How the PI 3-kinase/Akt pathway stimulates cell survival (Reprinted from Alberts et al. 2008, 934; with permission from Garland Science)

> 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. The entire network of pathways is depicted in Fig. 6.8.

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 in this case is the behavior of an animal cell to survive. The natural boundary of the object in question, the cell, is the cell membrane. The explanation of how a cell survives focuses on the processes that go on inside the cell. This is why Fig. 6.7 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.[57] But 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

---

[57] RTK is a transmembrane receptor protein. As such it is a part of the cell membrane and, hence, an internal factor.

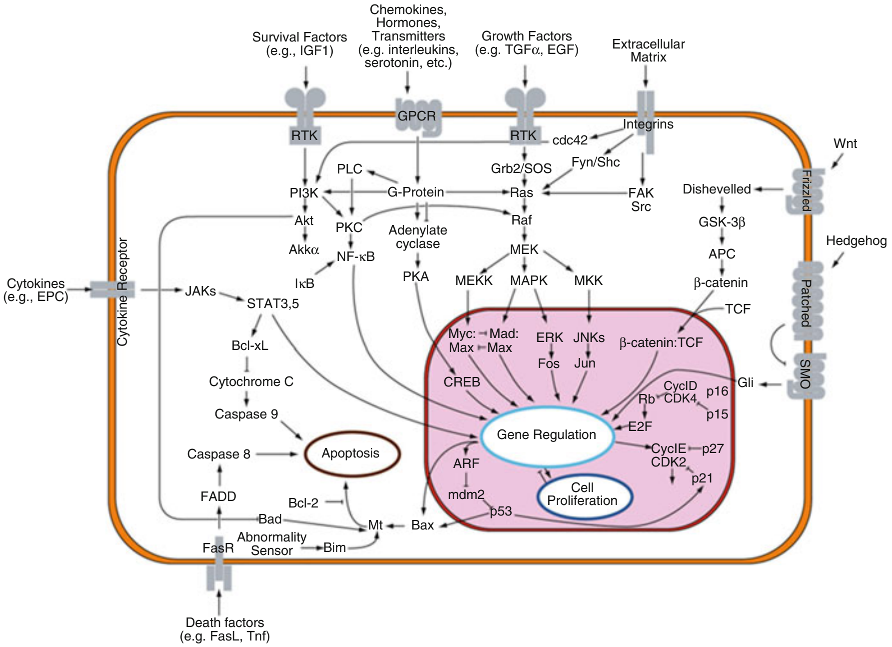

Fig. 6.8 Overview of signal transduction pathways involved in apoptosis

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 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.[58]

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 terms, they are classified as mere background conditions and, thus, ignored in the expla-

---

[58] 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)

nation. 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 biological object or system *Y* altogether, or to including selected parts of Y's environment, but simplifying them. The context of *Y* 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 and 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 study the parts of a biological object or system "in isolation" and to disregard how the parts are put together and integrated into the whole.[59] 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 in reductive explanations of a behavior of *Y* the biological parts of *Y* are represented as being *parts in isolation*. An important result of my analysis will

---

[59] 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).

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 limitations of reductionism (Sect. 4.1.1) I argue for how the phrase 'parts in isolation' is to be best understood (Sect. 4.1.2). My central claim will be that reductive explanations refer only to those relational properties of parts and to those kinds of interactions that can be discovered by studying the parts in other contexts that in situ. In Sect. 4.2 I use Bechtel's and Richardson's distinction of different kinds of decomposability (2010) to further explain this feature of reductive explanations. I examine whether the reductionist methodology can be characterized as treating biological objects as if they were aggregative or as if they were nearly decomposable.

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 an object, 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 and 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 and Hamilton 2007, 405–411). 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 an object (e.g. Dupré 1993, Chapter II, 2009). Others even identify reductive explanations with explanations, in which a biological object or 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 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 *neutral* regarding 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.[60]

---

[60] 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.

4.1 *Isolating Parts from Their Original Context*

What does it mean when biologists accuse reductionists of studying and explaining the behavior of a biological object or system *Y* by considering only *Y*'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; see also Kaiser 2011) 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 objects or systems.[61] This danger arises because to investigate the behavior of a biological object in a reductive manner implies to decompose the object into its biological parts and to try to achieve an understanding of the object as a whole by studying the behavior or properties of the parts *in isolation*. The following quotes exemplify this thesis[62]:

> The reductionist myth of simplicity leads its advocates to isolate parts as completely as possible and study these *isolated parts*. (Levins and 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 2004a, 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)

This gives rise to the question what is wrong with the reductionistic strategy to explore the parts of a biological object 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 biological object "in isolation" because, in order to understand the behavior of an object, we also need to know how these parts are integrated into the object

---

[61] "The reductionist approach… underestimates this complexity…" (van Regenmortel 2004a, 1016). "[R]eductionist thinking… strives to achieve a representation of reality free of complexity." (Soto and Sonnenschein 2010, 369)

[62] See also Greenspan 2001, 386; Kitano 2002, 1662; Chong and Ray 2002, 1661; and Strange 2005, 968.

as a whole.[63] 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 biological parts into a complete, functioning whole again (2010, xxxvii–xl). To recompose a biological object or system *Y* implies to show how the parts work together to bring about the behaviors that *Y* characteristically displays. Bechtel and Richardson describe this as disclosing the "orchestrated functioning" (2010, xxxix) of an object (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 and Chiriva-Internati 2005, 28; my emphasis)

This quotation exemplifies the major criticism biologists pass on reductionism: studying the behavior of a biological object *Y* in a reductionist manner, that is, by investigating *Y*'s biological parts in isolation, is defective because it does not throw light on how the individual parts are organized and how they "*dynamically interact*"[64] (Kitano 2002, 1662; my emphasis; see also Chong and Ray 2002; O'Malley and Dupré 2005, 1270; Sorger 2005, 9) in order to produce the behavior of the object 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)

---

[63] "[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 and Appenzeller 1999, 79; my emphasis; see also Bateson 2005, 35)

[64] The add-on 'dynamically' is superfluous since any interaction is temporally extended, involves changes and, as such, is dynamic and not static.

In Kitano's view, understanding a biological object or system (such as 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 object 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 biological object or system into parts, but also to study the parts in isolation; second, applying this reductionistic strategy results in the failure to capture how the biological parts are integrated into an object as a whole, that is, how they are organized and interact with each other.

4.1.2 Studying Parts in Different Contexts Than In Situ

On basis of this information I can now further specify what the phrase 'parts in isolation' means. One option is that the phrase 'studying parts in isolation' implies studying each biological part of an object *Y completely on its own*, that is, completely isolated from all other objects (including other parts of *Y*). According to this reading, the behavior or properties of a biological part of *Y* is explored under different conditions, but not in relation to other objects. This raises the problem that, by investigating 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 objects of certain kinds. Hence, if we understood the phrase '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 a biological part by isolating it completely from other objects is a research strategy that is alien to biological practice. Biological parts are not studied in complete isolation from anything else. Rather, in actual experimental settings the different conditions under which the biological parts of an object *Y* are studied usually include other objects (in particular other parts of *Y*).

Therefore, I favor an alternative interpretation, namely to understand the term 'parts in isolation' as 'parts isolated from their original context'. This means that the biological parts of an object or system *Y* are examined in a different context than in situ, for example, by taking them out of the original context and investigating them under laboratory conditions (i.e., in vitro), or by changing the original context for instance by inserting something into the object *Y* as a whole. These different contexts commonly include other objects, which is why the parts are not studied in complete isolation. The different contexts may even contain other biological parts of *Y*, but they need to be different from the context that surrounds the part when it is integrated with the other parts to the object or system *Y* 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 context) are still valid when the parts are put together again.[65] This assumption turns out to be particularly problematic in the case of so called integrated systems, 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 apply the gained insights to the concept of a reductive explanation. 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 and to focus on factors that are internal to the object *Y* as a whole. In addition, reductive explanations refer to *Y*'s biological parts as parts in isolation. But how are we to understand this thesis? What does it mean that an explanation represents the biological parts of *Y* as 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?

4.1.3 Referring Only to Some Kinds of Relational Properties and Interactions

I have already touched upon a related issue in Sects. 2.3.3 and 2.3.4. There the question was whether the lower-level character of reductive explanations precludes that functional properties and the organization of biological parts also are represented in the explanation. 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 biological parts of an object or system *Y* can be included in a lower-level explanation if these organizational features can be characterized as relational properties of lower-level objects (and not as properties of *Y*, i.e., of the whole). The question that is at issue in this section is similar to, but distinct from the one addressed in Sects. 2.3.3 and 2.3.4. 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 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 biological object 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 Chap. 4, Sect. 2.2). In the introduction to his volume on "Molecular Models of Life" (2005) Sarkar

---

[65] 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).

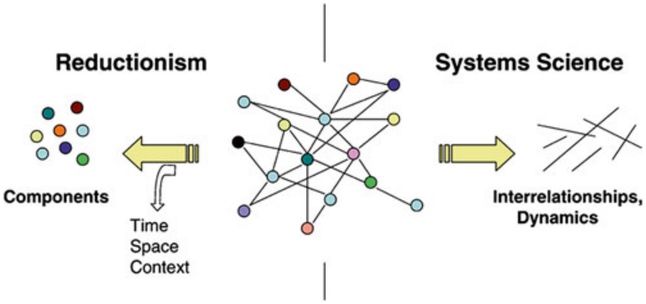

Fig. 6.9 Reductionism vs. systems science in medicine (Adapted from Ahn et al. 2006b, 1)

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 biological parts of an object or system *Y* since intrinsic properties are the only properties that a part has on its own, independent from the other parts of *Y*.[66] This implies that reductive explanations refer neither to the organization of *Y*'s parts, nor to their interactions because these are relational properties of the parts.[67]

This extreme view of reductive explanation and of reductionism in general can only rarely be found in the biological literature. A few biologists, however, endorse such a view. For instance, in a paper that applies central insights from systems biology to clinical medicine you can find Fig. 6.9. 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. 2006b, 1). That is, to investigate and explain the behavior of a biological object or system *Y* in a reductive way is identified with taking into account only the parts of *Y* with their intrinsic properties and ignoring the relations and interactions between the parts.

My impression is that, first, even if Sarkar's above quote suggests this, his 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

---

[66] By contrast, whether an object exhibits a relational (or extrinsic) property depends on the object's relationships to other things.

[67] 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).

reductive explanation (and of reductionism in general) in such a way that it applies only to explanations, which explain the behavior of a biological object or system *Y* by referring only to the intrinsic properties of *Y*'s biological parts.

Let me start with the first point. In Chap. 4, Sect. 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. An example illustrates this. Sarkar claims that even the phenomenon of cooperativity of hemoglobin, which traditionally has been regarded as a cornerstone of antireductionism (2005, 7f; van Regenmortel and Hull 2002, 254) can be explained reductively (2005, 7f, 10f, 110, 2008, 429). However, the explanation of cooperativity is far from appealing only to intrinsic properties of the biological parts of hemoglobin. Consider this example in more detail.

Example: Cooperative Binding of Hemoglobin (Physiology, Molecular Biology)

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.

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 in this case 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 biological 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 biological parts of hemoglobin. It also appeals to relational properties of the parts, such as 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. Hence, 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 biological parts of an object (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 object *Y excludes* all relational properties of and interactions between *Y*'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 biological parts of the object or system in question. 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 context'. That is, my central thesis is that to explain the behavior of a biological object *Y* by representing *Y*'s parts only as parts in isolation means to refer only to those relational properties of *Y*'s parts and to those kinds of interactions between them that can be discovered by investigating *Y*'s parts in other contexts than in situ. In other words, reductive explanations include only information that is discovered for example by taking the parts out of their original whole and studying them for instance in vitro, or by changing the original context for instance by inserting an alien object into the original whole and exploring the parts in this different context.

4.2 Treating Biological Systems as Nearly Decomposable

In the biological literature the concept of reductionism is not only associated with studying and explaining the behavior of a biological object or system *Y* by referring to *Y*'s parts as parts in isolation. In addition, two other specifications can be found: first, reductionism is treated as the view that biological objects are *aggregative systems*; second, a reductionist is said to take the organization of and interactions between the biological parts of an object 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 and to different kinds of decomposability that can be distinguished. I will argue that the first assumption (i.e., that a reductionist holds that biological objects are aggregative systems) yields not a convincing understanding of reductionism and reductive explanation, whereas the second does. Finally, I will integrate these results with the ones of Sect. 4.1 in order to identify a third major characteristic of reductive explanation in biology.

4.2.1 Aggregative Systems

Some biologists regard reductionism as the view that biological objects or 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 objects or systems are *aggregative systems* whose behaviors remain invariant to certain modifications of their biological parts. Nagel has termed this the “additive point of view” (1952, 26). The most extensive analysis of the notion of aggregativity has been developed by Wimsatt (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 biological object or system *Y* must remain invariant to certain modifications of *Y* (respectively, of *Y*’s parts).

Wimsatt’s Conditions of Aggregativity

(1) *Intersubstitution*: Invariance of Y’s behavior under interchanging its parts with parts from a relevant equivalence class.
(2) *Size scaling*: Qualitative similarity of the Y’s behavior under changes in the number of Y’s parts.
(3) *Decomposition and reaggregation*: Invariance of the Y’s behavior under decomposing and rearranging Y’s parts.
(4) *Linearity*: Not cooperative or inhibitory interactions among Y’s parts.[68]

The view that biological objects or systems are aggregative is problematic since in reality only very few of them 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. This view is underpinned by the fact that some biologists explicitly deny that reductionism can be identified with the view that biological objects or 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 struc-

---

[68] The forth condition reveals a close connection between the assumption that biological objects are aggregative systems and the assumption that the organization of and interactions between biological parts can be represented as neat, linear causal chains.

tures 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).[69]

4.2.2 Linear Causal Chains and Kinds of Decomposability

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 biological parts of an object or system, but who *underestimates their complexity* because he "analyses complex network interactions in terms of simple causal chains and mechanistic models" (van Regenmortel 2004a, 1018).[70] 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 consider the different kinds of decomposability of systems that can be distinguished.

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:

Kinds of Decomposability

(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

---

[69] 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)

[70] This underestimation of complexity is a part of what Dennett calls "*greedy reductionism*" (1995, 82).

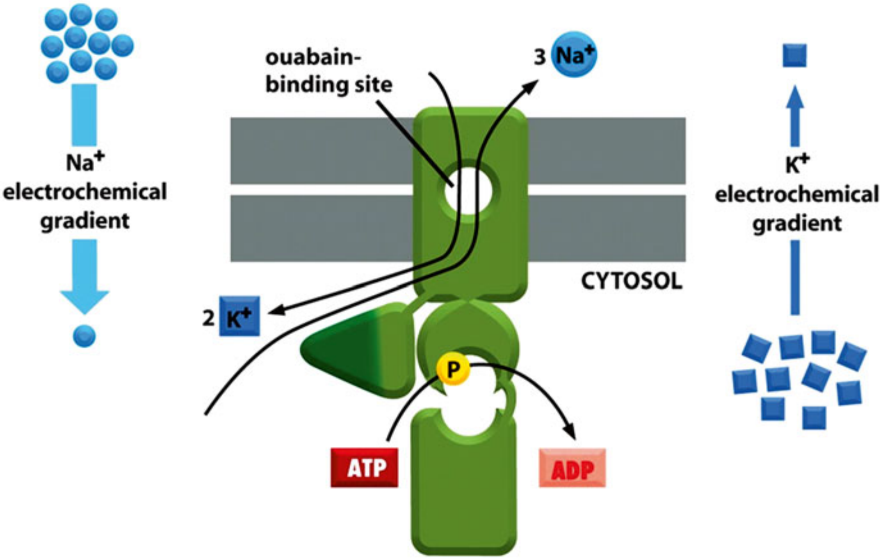

Fig. 6.10 The Na$^+$/K$^+$-ATPase (Reprinted from Alberts et al. 2008, 662; with permission from Garland Science)

The most extreme kind of a decomposable biological system is an aggregative system which satisfies Wimsatt's four conditions of aggregativity introduced in the previous section.[71] But Bechtel and Richardson qualify that "[f]ew interesting dynamic systems are strictly aggregative" (2010, 25).[72] As soon as the properties and behaviors of a biological object *Y* (the whole) are at least partially determined by how *Y*'s parts are organized, we no longer have aggregativity. To see this consider the example of the Na+/K+-ATPase.

Example: The Na+/K+-ATPase (Molecular Cell Biology)

The Na$^+$/K$^+$-ATPase is an ATP-driven antiporter that actively transports Na$^+$ ions out of the cell and pumps K$^+$ ions in against their electrochemical gradients (see Fig. 6.10). 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 Na$^+$/K$^+$-ATPase is far from being an aggregative system because it matters to the behavior of the Na$^+$/K$^+$-ATPase what its biological parts are, how they interact with each other (i.e., in which processes they are involved), and how they are spatially and temporally organized. The biological parts of the Na$^+$/K$^+$-ATPase might be intersubstitutable to a limited degree (since polypeptide chains with varying

---

[71] Bechtel and Richardson (2010) place special emphasis on Wimsatt's last condition of aggregativity since they are primarily concerned with exploring variations in organizational structure.

[72] Similarly, Wimsatt admits that "[i]t is rare indeed that all of these conditions are met" (1997, 375).

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)). The Na^+/K^+-ATPase thus 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 differ is the role that is played by the organization of their parts. 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 biological parts are more important for determining their properties than the interactions that occur between different parts. As a consequence, it is feasible to determine the properties of the parts of component systems *in isolation*.

Component systems are what Simon has called in his influential 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" (Simon 1962, 477).[73] This, however, does not mean that *no* interactions take place between the parts of component systems or that the parts are not organized. 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 nearly decomposable systems the contributions that the parts make to the behavior of the whole are sequential and linear, which is why the parts "retain an integrity of their own" (Bechtel and Richardson 2010, 199). As a result, organization and interactions provide only secondary constraints on the functioning of the biological parts (Bechtel and Richardson 2010, 26, 171).

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" and which contribute sequentially to the behavior of the whole? At first sight, this question can be answered affirmatively. As Fig. 6.11 illustrates, each of the biological parts of the Na^+/K^+-ATPase (e.g. its different subunits, ATP, Na^+, and K^+) is involved in a specific process that serves as the input to the next process. 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.

The properties of the biological 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

---

[73] See also Simon 1973. Wimsatt calls this kind of systems "interactionally simple" (1974, 72, 2007, 184).

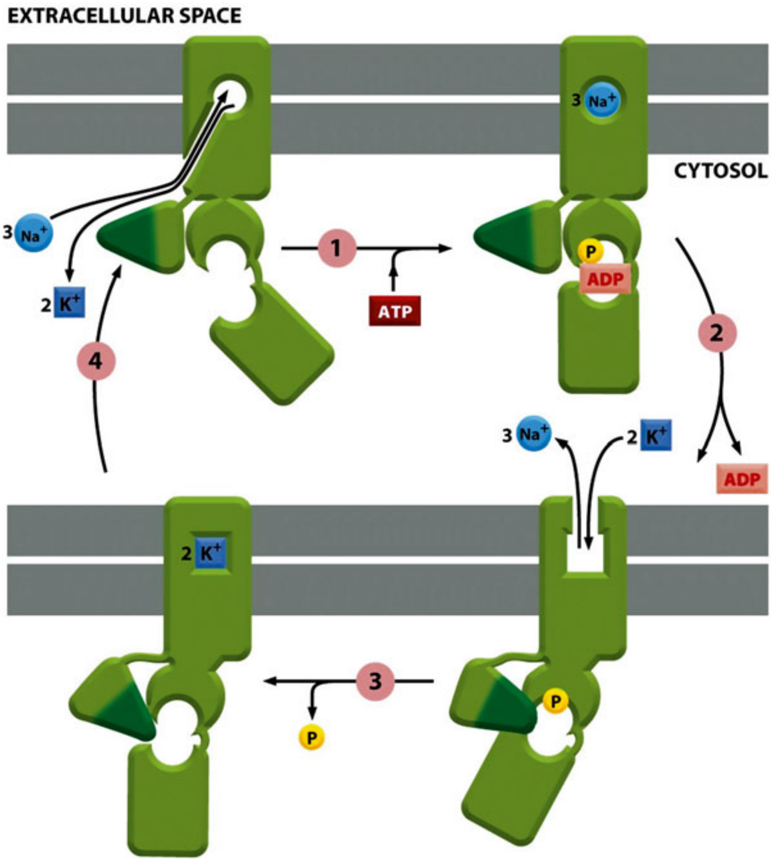

Fig. 6.11 The working of the Na+/K+-ATPase (Reprinted from Alberts et al. 2008, 662; with permission from Garland Science)

properties of the amino acid residues it is composed of. Accordingly, the properties of the biological parts of the Na+/K+-ATPase can be studied in isolation from their original context (e.g. in isolation from other parts of the Na+/K+-ATPase).

By contrast, in (functionally) *integrated systems* the properties of the parts are no longer intrinsically determined. Rather, how the biological parts are organized is said to be significantly involved in determining their properties and behaviors (2010, 26). This is the case because biological objects or systems of this kind exhibit a more complex form of organization (such as feedback relations and cyclic organization) and various interactions between their parts. This makes the behavior of the parts "mutually interdependent" (2010, 170) and restricts their autonomy.[74] According to

---

[74] 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)

Bechtel and Richardson, the biological parts of integrated systems have no "independent, isolable function" (2010, 31) and are thus only *minimally decomposable*.[75] Integrated systems are sometimes studied by assuming near decomposability. That is, they are dissected into isolated biological parts, which are assumed to make an independent, linear contribution to the behavior of the object or system as a whole. 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).

Finally, Bechtel and Richardson concede that there exist biological objects or systems, which are not decomposable into distinct biological parts at all and therefore decomposition appears to be "hopeless" or "misguided" (2010, 202) since the parts do not seem to perform intelligible subtasks contributing to the overall behavior of the system as a whole.[76] These cases are called *non-decomposable systems* since decomposition fails altogether.

Let us come back to the question from which we started: What does it mean to claim that a reductionist treats the interactions between the biological parts of an object or system as neat linear, causal chains? How is this related to representing biological parts as being 'parts in isolation'? I think Bechtel's and Richardson's distinction between nearly decomposable component systems and minimally decomposable integrated systems is of help here. My proposal is to link reductionism to the view that biological objects or 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 an object or system in a reductive manner means to treat them as if they were composed of biological parts, whose properties are mainly unaffected by their context and which, thus, can be studied in isolation from their original context (i.e., in other contexts than in situ). Furthermore, to reductively explore a biological object *Y* implies to assume that the organization of *Y*'s biological parts 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 an object or 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 biological object or system *Y* of concern is an integrated rather than a component system, it may result in serious errors to study *Y*'s biological parts in isolation and to assume that the parts make an independent, linear contribution to the behavior of *Y* (the whole). 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 how they are organized. Consequently, biological parts often exhibit different properties (or are involved in

---

[75] This high degree of integration can be traced back to the fact that the component subsystems have evolved together (Levins 1970, 77).

[76] Bechtel and Richardson primarily discuss "connectionist systems" as examples for non-decomposable systems (2010, 199–229).

different processes) when they are integrated into the object or system as a whole (i.e., in situ) and when they are placed in different contexts than the original one. 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 (recall Sect. 2) and focus on those factors that are internal to the biological object or system *Y* in question (recall Sect. 3), they also appeal to *Y*'s parts as if they were parts in isolation. The latter characteristic means that reductive explanations describe only those relational properties of and interactions between *Y*'s biological parts that can be discovered by studying the parts in other contexts than in situ. This implies treating the biological object or system of concern as a nearly decomposable system, that is, as an object or system whose biological 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

In the previous three sections I developed the core of my account of explanatory reduction which can be summarized as follows. Reductive explanations exhibit two necessary features, (1) and (3), and one feature (2), which is typical but not necessary.

My Account of Explanatory Reduction

Reductive explanations in the biological sciences explain the behavior of a biological object or system *Y* by

(1) referring only to factors that are *located on a lower level* of organization than *Y*,
(2) focusing on factors that are *internal to Y*, that is, that are biological parts of *Y*, and
(3) representing the biological parts of *Y* as if they were *parts in isolation* (i.e., representing only those relational properties of and interactions between the parts that can be studied in other contexts than in situ).

There are many reasons in favor of the view that, taken together, these three conditions are sufficient for a biological explanation to be reductive.[77]

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 biological explanation can be specified.

5.1 Part-Whole Explanation

Part-whole explanations are paradigmatic cases of reductive explanations, which is why some philosophers identify reductive explanations with part-whole explanations (e.g. Hüttemann and 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 Sect. 3 I pointed out that all explanations, which explain the behavior of a biological object *Y* (the whole) by appealing only to factors that are internal to *Y*, that is, that are biological parts of *Y*, are part-whole explanations. However, I also argued that for an explanation to be reductive it need not invoke only the biological parts of *Y*. It may also refer to environmental factors (or external factors) provided that these factors are located on a lower level of organization than the object *Y* whose behavior is to be explained. In short, the internal character is *not* a *necessary condition* for the reductivity of an explanation, whereas the lower-level character is. The claim that the lower-level character and the internal character of an explanation can fall apart requires the view that levels are not only determined by particular part-whole relations, but also by hierarchies of part-whole relations and by the membership of objects to certain general biological kinds (Kaiser manuscript a; recall Sects. 1.2 and 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* (as they include external factors without simplifying them). 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 which I discussed in Sect. 4. Not all part-whole explanations explain the behavior of the whole *Y* by appealing only to those relational properties of and interactions between *Y*'s biological parts that can be discovered by studying the parts in isolation (i.e., by treating *Y* as a nearly decomposable system). This is exactly what biologists who call for a move "beyond reductionism" (Gallagher and Appenzeller 1999, 79) are driving at: the behavior especially of complex systems cannot be adequately explained by taking into account only parts in isolation. Rather,

---

[77] 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.

Table 6.1 Comparing reductive and part-whole explanations

	Reductive explanation	Part-whole explanation
Lower-level character	necessary	necessary
Internal character	not necessary	necessary
Parts-in-isolation character	necessary	not necessary

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 biological object or 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).[78] Table 6.1 sums up the relation between reductive explanation and part-whole explanation.

5.2 *Mechanistic Explanation*

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)

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 and Chiriva-Internati 2005, 28)[79]

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 and Millstein 2005, 327) in contemporary philosophy of science. Rather, their understanding of 'mechanism' traces back to the Mechanistic Philosophy of the late sixteenth and seventeenth century and to the subsequent dispute about mechanicism and vitalism in the eighteenth and nineteenth century. This also makes clear why they relate the notion of mechanism so closely to the notion of reduc-

---

[78] 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.

[79] See also van Regenmortel 2004a, 1018; O'Malley and Dupré 2005, 1270; and Bizzarri et al. 2008, 184.

tion. In the Mechanistic Philosophy of the late sixteenth and seventeenth century René Descartes (1991 [1644]), Robert Boyle (1966), and others argued that all natural phenomena can be made intelligible by mechanistically explaining them in terms of material particles, their few, simple properties (such as size, shape, and motion), and the interactions between them.[80] 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 what I call *fundamental-level explanations* – with the fundamental level being not the lowest biological level of molecules, but rather the lowest physical level that was assumed at that time, that is, the level of material particles and of their properties size, shape, and motion.

In contemporary philosophy of science the Mechanistic Philosophy experiences its second spring. Since the sixteenth and seventeenth the century, however, 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 objects and their properties as well as many different forms of activities, interactions, and processes. 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. In these days the notion of mechanistic explanation is thus not confined to the class of fundamental-level explanations anymore (with the fundamental level being the level of material particles with their properties size, shape, and motion).

Before I discuss whether mechanistic explanations exhibit the three characteristics that distinguish reductive explanations in biology, 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 and Richardson 2010, xiii, xvii, xxxvii–xl). On the other hand, 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 (recall Chap. 3, Sects. 3.1 and 4). They do not oppose the idea of reduction in general.[81] The New Mechanists thus appear to maintain the old alliance between mechanisms and reduction. But how far does this

---

[80] See for instance Boyle, "The Excellency of Theology" (1966), 69–71, 77f and Descartes, "Principles" (1991 [1644]), 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 and Chiriva-Internati 2006, 5; my emphasis).

[81] This is apparent, for instance, in Machamer et al. 2000, 23; Bechtel and Richardson 2010, xxxvii; and Bechtel and Hamilton 2007, 405.

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 the biological sciences? In what follows I successively examine whether mechanistic explanations exhibit the three characteristics of reductive explanations.

First, I have just argued that mechanistic explanations cannot be identified with fundamental-level explanations. But are they lower-level explanations? That is, are all the components of a certain mechanism located on a lower level of organization as the mechanism as a whole? At first sight it seems as if also mechanistic explanations span (at least) two levels, namely the level of the mechanism as a whole and the level of its components. But remember that according to my view of levels (Kaiser manuscript a), not *component-mechanism relations* but *part-whole relations* in general (co-)determine levels and that the former differ in important respects from the latter (recall Sect. 1.2): first, the components of a biological mechanism are individuated with regard to the one behavior the mechanism is for, whereas the biological parts of an object are individuated with regard to all of the characteristic behaviors of the object; second, biological mechanisms often transgress the natural boundaries that constrain biological objects and thus include objects (and their processes) as components that are not part of the corresponding biological object in question (cf. Kaiser and Krickel forthcoming). These differences have important implications for whether mechanistic explanations possess a lower-level character or an internal character.

Consider the lower-level character of reductive explanations first. In most cases the components of a mechanism for the behavior of a certain biological object *Y* will also be biological parts of the object *Y* and thus be located on a lower level than *Y*. However, in some cases a mechanism will include objects (and the processes in which they are involved) that are external to *Y* (i.e., not parts of *Y*). It might be that these external objects belong to one of the general biological kinds to which the biological parts of *Y* belong (in which case they would also be located on a lower level than Y, recall Sect. 1.2). But these external objects might also be higher-level factors. Even if such examples are rare they are clear cases in which mechanistic explanations and reductive explanations fall apart because the mechanistic explanation does not exhibit a lower-level character.

Second, consider now the second feature of reductive explanations, their internal character. The first difference between component-mechanism relations and part-whole relations in general that I have just pointed out implies that the set of all components of the mechanism for a behavior of Y is smaller than the set of all biological parts of object Y (since parts are individuated with regard to all characteristic behaviors of Y, not only with regard to the one behavior the mechanism in question is for). Thus, there exist reductive explanations that exhibit an internal character, but that are not mechanistic explanations.

In turn, it is also questionable whether mechanistic explanations must possess an internal character because mechanisms are said to often transgress natural boundaries and refer to components that are not biological parts of the corresponding bio-

Table 6.2 Comparing reductive and mechanistic explanations

	Reductive explanation	Mechanistic explanation
Lower-level character	necessary	not necessary
Internal character	not necessary	necessary
Parts-in-isolation character	necessary	not necessary

logical object. The mechanism for muscle contraction, for example, includes neurotransmitter molecules that are located outside of the cell membrane of the muscle fiber and thus are not biological parts of the muscle fiber. In the majority of cases, however, these components of mechanisms that are external to the corresponding biological object are simplified as being input or background conditions,[82] which is why the internal character of the mechanistic explanation is preserved. The internal character thus seem to be 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 biological object *Y* by reference to Y's parts in isolation is characterized as analyzing "complex network interactions in terms of simple causal chains and mechanistic models" (van Regenmortel 2004a, 1018). In accordance with that, some mechanists focus in their work on mechanisms with relatively simple organizations (e.g. Glennan 2002, 2008; Tabery 2004). But the majority of the New Mechanists argue that their account equally applies to more complex kinds of biological objects and systems such as integrated systems that exhibit complex forms of organization like feedback (see, in particular, Bechtel and Richardson 2010; Bechtel 2001, 2009; Bechtel and Abrahamsen 2011). They thus 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 a mechanism in isolation (i.e., in other contexts than in situ). The result of this focus on complex phenomena and systems is that the set of mechanistic explanations contains not only reductive explanations but also encompasses non-reductive explanations that may possess a lower-level and an internal character, but that do not meet the third condition of reductivity, that is, they do not appeal to the mechanistic components only as parts in isolation. Table 6.2 sums up the relation between reductive explanation and mechanistic explanation.

---

[82] The New Mechanists claim that "set-up conditions", "enabling conditions", or even background conditions, which often include external factors, should be regarded as "parts of the mechanism" (Machamer et al. 2000, 11).

6 The Ontic Character of My Account

The analysis of what makes a biological explanation reductive that I have developed in this chapter can be characterized as an *ontic* account of explanatory reduction. But what is it that makes my account ontic? In Chap. 5, Sect. 2 I made the first steps towards answering this question. I pointed out that, in the philosophy of science, the term 'ontic' is commonly used to denote a certain kind of view about scientific explanation. In this debate, ontic conceptions of explanation are contrasted with epistemic conceptions (originally, this distinction was tripartite, but contemporary discussions often ignore Salmon's modal conceptions of explanation; cf. Salmon 1984a).

My main claim in Chap. 5, Sect. 2 was that two equally legitimate views exist about what makes a conception of explanation ontic and what distinguishes ontic from epistemic accounts of explanation. According to the *strong reading*, the ontic character of a conception of explanation traces back to the claim that explanations exist in the world, independently from scientists discovering them, from inquirers requesting them, or from speakers uttering them. Those who adopt the strong reading regard the ontic-epistemic dispute as a dispute about the kind of entity an explanation is: Are explanations out there in the world, as the ontic account states, or are explanations parts of our knowledge about this world, are they representations or descriptions of this world, as the epistemic account asserts? According to the *weak reading*, both ontic and epistemic accounts agree that explanations are epistemic entities. What distinguishes them are their different claims about where the explanatory force of an explanation stems from: ontic accounts trace explanatory power back to the fact that certain objects and relations in the world are truly represented, whereas epistemic accounts trace explanatory power, for instance, back to certain logical relations that hold between sentences constituting the explanation. After presenting these two different interpretations of the ontic character of an explanation I argued that the strong reading faces some serious problems, which is why I use the term 'ontic' analogous to the weak reading when I characterize my own account of explanatory reduction as ontic.

So far so good, but how can this analogy between ontic conceptions of explanation (in their weak reading) and ontic accounts of explanatory reduction be spelled out in detail? What makes my analysis of reductive explanation in biology ontic? In a nutshell, my account of explanatory reduction is ontic 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 that exist in the world. Contrary to proponents of a Nagelian model of theory reduction (recall Chap. 3, Sects. 3.1 and 4), I regard formal issues (such as 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 the formal but on what Sarkar calls the "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 biological object or system *Y* is explained by, first, appealing only to factors that are located on a *lower level* of organization than *Y*, second, by focusing on factors that are *internal* to *Y*, that is, that are biological parts of *Y*, and third, by describing Y's biological parts as if they were *parts in isolation* (i.e., by describing them as objects, whose properties are not determined by how they are organized 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 my view, what determines the reductivity of an explanation is whether they truly represent certain relations that exist in the world. I identify three of these relations: first, the relation that an object is located on a higher, on a lower, or on the same *level of organization* than another object; second, the relation that an object is internal to another object (the whole), that is, that it is a *biological part of* the whole; and third, the fact that an object consists of *parts in isolation*, that is, of biological parts whose properties are not determined by their organization and that interact with each other in a linear, sequential way. If an explanation satisfies the above conditions, but is not true (e.g. because the properties of the biological parts 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 Chap. 3, Sect. 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 biological object and it's organized, interacting biological 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[83] that exist in the world.[84]

---

[83] 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 can but need not be characterized as instances of ontological reduction (recall Chap. 3, Sect. 2.1.1).

[84] This, however, does not imply that the world completely determines the reductivity of an explanation. As I point out in Kaiser (forthcoming a), pragmatic factors might have an influence on how the line between a biological object or system and its environment is drawn and on which objects and processes are identified as biological parts. In this way pragmatic factors might also affect the reductive character of an explanation.

7 Interim Conclusion

In this chapter, I developed my own account of explanatory reduction in biology by analyzing paradigmatic and important examples of reductive explanations from biological practice as well as reasoning strategies that biologists employ when discussing reductionism. The central question that I answered was: Which characteristics are common to reductive explanations in the biological sciences and distinguish them from non-reductive explanations? I argued that reductive explanations exhibit two necessary features, (1) and (3), and one feature (2), which is typical but not necessary.

My Account of Explanatory Reduction

Reductive explanations in the biological sciences explain the behavior of a biological object or system *Y* by

(1) referring only to factors that are *located on a lower level* of organization than *Y*,
(2) focusing on factors that are *internal to Y*, that is, that are biological parts of *Y*, and
(3) representing the biological parts of *Y* as if they were *parts in isolation* (i.e., representing only those relational properties of and interactions between the parts that can be studied in other contexts than in situ).

I claim 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.

This account of explanatory reduction is ontic because it traces back the difference between reductive explanations and non-reductive explanations to certain kinds of relations that exist in the world, rather than to certain logical relations between explanandum-sentence and explanans-sentence. In my view, what determines whether an explanation is reductive or not is whether or not it truly represents specific kinds of relations that exist in the world, namely, the relations 'being located on a lower level of organization than', 'being internal to' (i.e., 'being a biological part of'), and 'being parts in isolation' (i.e., 'being nearly decomposable').

This view raises the question as to what the nature of these relations is and under which conditions they exist. I assume that the existence of a part-whole relation between two biological objects *X* and *Y* requires that three necessary conditions are fulfilled: first, X must be spatially included in *Y* (i.e., if *Y* has a natural boundary then *X* must be spatially located inside or in the region that *Y*'s natural boundary occupies), *X* must be involved in biological process *A* and *Y* must characteristically exhibit behaviors $B_1, \ldots B_n$ so that, second, *A* is temporally included in $B_1, \ldots B_n$ (i.e., *A* takes place during any of the time periods in which *Y* exhibits $B_1, \ldots B_n$), and third, *A* is relevant to at least one of *Y*'s behaviors $B_1, \ldots B_n$ (the details are worked out in Kaiser forthcoming a). My conception of different-level relations is based on this notion of biological parthood. In my view, two major factors determine whether a biological object *X* is located on a lower level than biological object *Y*: Either *X* is a biological part of *Y* or *X* belongs to a general biological kind to which at least one of *Y*'s biological parts belong (cf. Kaiser manuscript a). This notion of level enabled me to clarify why lower-level factors are more than parts and why the two corresponding features of reductive explanation, their lower-level character and their internal character, cannot be lumped together. The set of lower-level factors does not only contain the biological parts of the object or system *Y* under consideration; it comprises also such factors that are external to *Y* but are located on a lower level than *Y* because they belong to the same general biological kind as *Y*'s parts. Hence, in biological explanations, the internal character can be violated while the lower-level character is retained (in the case of which, the resulting explanation still may be reductive).

Philosophers have not yet paid sufficient attention to the third kind of relation that determines the reductivity of a biological explanation. It is the relation of 'being a part in isolation'. To be a part in isolation means to be a biological part that satisfies further, additional conditions: one that interacts with other parts of the same whole in a sequential and linear fashion (i.e., the whole is organized in a simple way) and exhibits properties that are largely unaffected by their context. Biological systems that are, in Simon's and Wimsatt's terms, nearly decomposable consist of such parts in isolation. To be clear, my claim that reductive explanations in biology refer to parts only as parts in isolation does not imply that all biological objects or systems *in fact* are nearly decomposable and simply organized. They are *not*, and this is exactly the reason why reductionism has its limits (Kaiser 2011), why the parts of some biological systems cannot be studied in other contexts than in situ, why the search for reductive explanations sometimes fails or results in inadequate explanations. If the behavior of a biological system is explained by describing its parts in isolation but the system is an integrated system, the resulting explanation will be reductive but inadequate.

My analysis of reductive explanation in biology yields further central insights: First, I clarified what exactly it means for a biological explanation to refer solely to lower-level factors and to exclude higher-level factors. I pointed out that reference to functional properties of lower-level objects does not violate the lower-level character of an explanation as long as the context that needs to be included is also spelled out in lower-level terms. Relatedly, I argued that the organization of *Y*'s parts counts as a lower-level factor if it can be characterized as a lower-level (relational) property of a part of *Y* and not as a property of *Y*, the whole, itself. Second, I introduced two significant subtypes of lower-level explanations, fundamental-level explanations (i.e., molecular explanations) and single-factor explanations (which rarely occur in biological practice). I claimed that, although molecular explanations are an important subtype of reductive explanations, philosophers should refrain from identifying reductive explanations with fundamental-level explanations because this identification obscures the diversity of reductive explanation in the biological sciences and converts explanatory reductionism into a straw man. Third, reductive explanations focus on internal factors in the sense that either they ignore environmental factors altogether or they simplify them to a great extent (i.e., conceptualize them as background conditions or as mere input conditions). Finally, my account of explanatory reduction can be used 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 Table 6.3.

Table 6.3 Comparing reductive, part-whole, and mechanistic explanations

	Reductive explanation	Part-whole explanation	Mechanistic explanation
Lower-level character	necessary	necessary	not necessary
Internal character	not necessary	necessary	necessary
Parts-in-isolation character	necessary	not necessary	not necessary

Chapter 7
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 is to provide an understanding of an important element of contemporary biological research practice, namely, of *reductive explanation*. My main question is: 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 Chap. 6, I answer this question by developing my own account of explanatory reduction in biology. The other four chapters, however, do more than merely setting up the discussion of the rest of the book. They also yield novel insights, the most important one of which I will summarize in the following.

The primary goal of my meta-philosophical analysis in Chap. 2 is to disclose the aim, methodology, and criteria of adequacy of my analysis of reductive explanation in biology. That way, it also contributes to clarifying how philosophy of science in general can or should be pursued and which criteria of adequacy a philosopher of science in practice should accept. One important result of my analysis is that, in the philosophy of science, different kinds of projects are pursued, which can be arranged on a spectrum depicted in Fig. 7.1. The most important kinds of projects are the three in the middle of this spectrum. My own project of developing an account of explanatory reduction in biology is a project of the descriptive-critical kind. Normative projects are distinct from but equally legitimate as descriptive-critical projects. 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.

This classification allows me to distinguish three different dimensions of normativity that occur in the philosophy of science:

M.I. Kaiser, *Reductive Explanation in the Biological Sciences*, History, Philosophy and Theory of the Life Sciences 16,
DOI 10.1007/978-3-319-25310-7_7

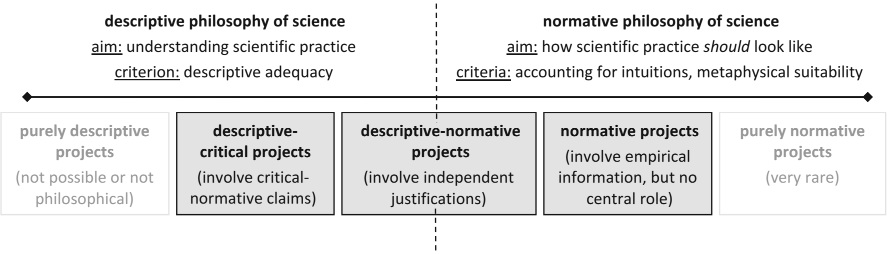

Fig. 7.1 A spectrum of different kinds of projects in philosophy of science

Dimensions of Normativity in the Philosophy of Science
A philosophical account of a certain element of science X is normative if

(1) **Methodological normativity** (descriptive projects)

- it contains factual claims about X and
- it (implicitly or explicitly) relies on *normative claims about how to select the empirical information* about X on the basis of which the account is developed. Or

(2) **Normativity in the strict sense**

- it contains *normative claims about* X (and)
- [it contains factual claims about X, but they are not crucial for the development and justification of the account (only for normative projects).] Or

(3) **Norm-normativity**

(a) it contains factual claims about X (where X is a social or epistemic norm) or about how X relates to epistemic norms (*description of norms*) or
(b) it contains normative claims about X (where X is a social or epistemic norm) or about how X relates to epistemic norms (*justification of norms*).

My own account of explanatory reduction involves methodological normativity (dimension 1) and norm-normativity as description of norms (dimension 3a). Methodological normativity is quite widespread in the philosophy of science (in practice). It makes explicit the fact that philosophical approaches cannot be purely descriptive but must involve a critical-normative stance on the empirical data that scientific practice provides them with. In my view, 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. My analysis involves also is norm-normativity because I specify not only what reductive explanations in biology are but also what their relations are to certain epistemic norms such as explanatory success. But since my account does not include the independent justification of epistemic norms such as explanatory success it displays norm-normativity only as a critical description or reconstruction of those norms that are in fact accepted in biological practice.

Another major result of my meta-philosophical analysis was a list of four criteria that determine the adequacy of my account of reduction: descriptive adequacy, balancing specificity against generality, norm-normativity, and potential usefulness. At the end of this conclusion I will get back to these criteria and show how my account satisfies them.

In Chap. 3, I introduce the previous debate about reductionism in the philosophy of biology by pointing out four lessons that one should learn from it. These lessons are the following: First, before you discuss whether reductionism or antireductionism is true you should seek to understand what reduction is. 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. 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 independent. 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. 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 biological practice.

Besides these four lessons I want to emphasize two further crucial results of Chap. 3. The one concerns the notion of ontological reduction and the other the concept of a reductive method. First, I specified the notion of ontological reduction(ism) that is, explicitly or implicitly, taken for granted in the philosophy of biology. My analysis discloses 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). Instead, 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. Furthermore, I pointed out 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 is the difference between constitution and identity 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, I directed attention to the fact that in the current philosophical literature on methodological reductionism it is left unclear what exactly it means to pursue biological research in a reductive fashion. I try 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 is continued in Chap. 6. 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 biological object or system of interest (which implies conceptualizing contextual factors as background or as input conditions) and to study the parts of a biological object or system in isolation (which I specified as studying the parts in other contexts than in situ).

In Chap. 4, I critically discuss the previous work that has been carried out on the topic of explanatory reduction in biology. I examine 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, and Sarkar's, Hüttemann's, and Love's work on the reductive character of individual biological explanations. The goal of this critical discussion is to identify fruitful insights that can be utilized in developing my own account and to sort out those ideas about explanatory reduction that prove to be unconvincing. I conclude that Rosenberg's perspective encounters devastating criticism. 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.

In line with this choice, I argue 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 arises from this perspective is: Under which conditions does the relation between explanandum and explanans count as a reduction? In other words, which characteristics of biological explanations determine their reductive character?

Chapter 5 reveals the interrelations between the issue of explanation and my account of explanatory reduction. I also addressed questions about explanation that are involved in disputes about explanatory reductionism but only insofar as they affect my analysis of reductive explanation. Three results of this chapter are of particular importance. First, there are two equally legitimate readings of the ontic conception of explanation. According to the strong reading, what makes an account of explanation ontic is the thesis that explanations are objects or facts that exist in the world itself (rather than epistemic units like representations, models, or texts). By contrast, according to the weak reading, 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. What speaks in favor of the weak reading is that strong readings of the ontic conception of explanation face serious objections. This is why I rely on the weak reading of the term ‘ontic’ when characterize my own account of explanatory reduction as ontic.

Second, the question what constitutes the reductive character of biological explanations (i.e., the question of explanatory reduction) does not boil down to the question what characterizes an adequate explanation (i.e., the question of explanation). What makes an explanation reductive is different from what makes a description explanatory. In line with this, I argue 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. This neutrality is a virtue rather than a shortcoming of my account.

Third, 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). 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 which, 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. Hence, in my view, explanation inevitably has pragmatic dimensions. This does, however, not imply that the adequacy of an explanation is “subjective” or exclusively determined by pragmatics. 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.

Chapter 6 contains my own analysis of the characteristics of reductive explanations and of the merits and limitations of explaining biological phenomena in a reductive manner. In a preliminary section I introduced two concepts that occupy center stage in my account of explanatory reduction: the concept of a biological part and the concept of a level of organization. The main thesis of the account of biological parthood that I develop in Kaiser (forthcoming a) is that the existence of part-whole relations in the biological realm requires three conditions to be satisfied: spatial inclusion, temporal inclusion, and relevance. I specify these three conditions as follows:

Biological Parthood
Object *X* is a biological part of object *Y* iff

(1) if *Y* has a natural boundary then *X* is *spatially located inside* or *in* the region of the natural boundary (**Spatial Inclusion**),
(2) *X* is involved in biological process *A*, *Y* characteristically exhibits behaviors B_1,… B_n, and *A takes place during* any of the time periods in which *Y* exhibits B_1, B_2, or… B_n (**Temporal Inclusion**), and
(3) *A is relevant* to at least one of the behaviors B_1,… B_n that *Y* characteristically exhibits (**Relevance**).

I take levels to be determined by part-whole relations in general, not by relations between mechanisms as wholes and their components. The major shortcoming of a mechanistic account of levels is that it results in a too restricted view of levels, according to which levels are only determined in local explanatory contexts. Such a local notion of levels does not capture level claims that biologists commonly make in their research practice. In addition, a local notion of levels does not meet the demands of my analysis of reductive explanation. I identify a second factor that determines levels, namely, the membership of biological objects to general biological kinds. This factor makes the account more global, without assuming the implausible view that the natural world is divided into a few, monolithic levels. In sum, my account of levels is the following (Kaiser manuscript a):

Levels of Organization
Object *X* is located on a lower level than object *Y*

(1) *X* is a *biological part* of *Y* (**Biological Parthood**) *or*
(2) *X* belongs to the same *general biological kind* as one or more of the biological parts of *Y* (**Kind Membership**).

Based on these conceptual clarifications I could then address the central question of my book, namely which characteristics are common to reductive explanations in the biological science, and allow distinguishing them from non-reductive explanations. The core thesis of my account of explanatory reduction is that reductive explanations exhibit two necessary features, (1) and (3), and one feature that many but not necessarily all reductive explanations exhibit, (2). Taken together, these three conditions plausibly are sufficient for a biological explanation to be reductive.

My Account of Explanatory Reduction

Reductive explanations in the biological sciences explain the behavior of a biological object or system *Y* by

(1) referring only to factors that are *located on a lower level* of organization than *Y*,
(2) focusing on factors that are *internal to Y*, that is, that are biological parts of *Y*, and
(3) representing the biological parts of *Y* as if they were *parts in isolation* (i.e., representing only those relational properties of and interactions between the parts that can be studied in other contexts than in situ).

I elaborate on these three characteristics of reductive explanations in turn, starting with their lower-level character. Reductive explanations in the biological sciences 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). But what, exactly, does it means for an explanation to refer solely to lower-level factors and to exclude factors from higher levels? My analysis yields three main results: first, the set of lower-level factors is not confined to the set of biological parts of a particular biological object *Y*. Rather, it comprises also factors that are external to *Y*, but that are located on the same level(s) as the biological parts of *Y*. Second, the category of lower-level factors also encompasses functional properties of biological parts. 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 also described in lower-level terms. Third, the organization of *Y*'s biological 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 property of the biological object as a whole.

Another major result of my analysis is the claim that there are two subtypes of lower-level explanations, namely fundamental-level explanations and single-factor explanations. Fundamental-level explanations are reductive explanations that refer only to factors that are located on the lowest, the fundamental biological level of molecules. Hence, they are molecular explanations. I argue that even though fundamental-level explanations constitute an important subtype of reductive explanations, philosophers should not be fooled into identifying reductive explanation with fundamental-level explanation. Doing so obscures the diversity of reductive explanation in biological science and converts explanatory reductionism into a straw man. Single-factor explanations explain the behavior of a biological object by referring only to a single causal factor. 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).

A second feature that is typical (but not necessary) for the reductivity of biological explanations is their internal character. Many reductive explanations explain the behavior of a biological object or system *Y* by focusing on those factors that are internal to *Y*. A factor is internal to *Y* iff it is a biological part of *Y*. Hence, explanations with an internal character are part-whole explanations. Because the internal character of reductive explanations has been overlooked or intermingled with their lower-level character for a long time it is important to be aware of the differences between these two features of reductive explanations. My analysis shows that the category of lower-level factors is broader than the category of internal factors. That is, all explanations with an internal character are lower-level explanations, but the set of lower-level explanations also includes reductive explanations that fail to possess an internal character.

The second condition of my account states that (most) reductive explanations focus on factors that are internal to Y (i.e., that are biological parts of Y). The term 'focus on' expresses that either reductive explanations ignore environmental factors altogether (i.e., they do not mention them at all) or they simplify environmental factors to a great extent. I argue that the context of a biological object or system can be simplified in two major ways: environmental factors can be simplified as background conditions, which are treated as being homogeneous or remaining constant over time, or 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.

My thesis that the internal character is a typical but not a necessary condition for the reductivity of a biological 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 also be reductive. 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 (cf., e.g., Weber 1998). 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 object inevitably appeal to contextual factors, they can still be reductive if these contextual factors are described in lower-level terms.

A third characteristic of reductive explanations in biology is that they explain the behavior of a biological object or system *Y* by appealing to *Y*'s biological parts as parts in isolation. Philosophers have not paid much attention to this necessary feature of reductive explanations yet but the current vivid discussions in biology about the limits of reductionism show that they should. I argue that the phrase 'to study parts in isolation' does not mean that the biological parts of an object are investigated completely on their own but rather that the parts are studied in isolation from their original context (i.e., in other contexts than in situ). Studying parts in isolation does not result in reductive explanations that refer only to intrinsic properties of *Y*'s parts. Rather, 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 than in situ. This third feature of reductive explanations can be rephrased as describing the biological parts of an object or 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 other words, reductive explanations describe the biological system to be explained as a nearly decomposable system. This third feature of reductive explanations reveals why reductionism has its limits. If a biological object or system *Y* is *not* nearly decomposable (but, e.g., only minimally decomposable) representing *Y*'s parts as parts in isolation and treating *Y* as a nearly decomposable system will probably result in a reductive explanation, but in one that is *inadequate* because it misrepresents the properties of and interactions between *Y*'s parts.

My account of explanatory reduction also enables me to clarify the similarities and differences between three kinds of explanation, namely, reductive explanation, part-whole explanation and mechanistic explanation. The main results are summarized in Table 7.1.

In this book, the concept of a mechanistic explanation constituted only a side issue, which is why many interesting questions remained undiscussed. In particular, the similarities and differences between mechanistic explanation and part-whole explanation (and between the underlying component-mechanism relations and part-whole relations) are promising objects of future research (see, e.g., Kaiser and Krickel forthcoming).

Table 7.1 Comparing reductive, part-whole, and mechanistic explanations

	Reductive explanation	Part-whole explanation	Mechanistic explanation
Lower-level character	necessary	necessary	not necessary
Internal character	not necessary	necessary	necessary
Parts-in-isolation character	necessary	not necessary	not necessary

I characterize my account of explanatory reduction in biology as an ontic account. The term 'ontic' 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. According to my account, what determines whether an explanation is reductive is whether it truly represents the explanandum phenomenon as possessing specific relations to the factors cited as explanatorily relevant. My account identifies three of these relations in the world: 'being located on a lower level of organization than', 'being internal to/a biological part of', and 'being a part in isolation' (i.e., 'being a part whose properties are not determined by the organization of the whole *Y* and that interacts with other parts of *Y* in a linear, sequential way'). It must be noted that this understanding of the term 'ontic' is different from the one that for instance Craver and Strevens assume. Contrary to their view, I stick to the claim that (reductive) explanations are representations of entities in the world rather than, themselves, being entities in the world. This weak reading of the ontic character of an account avoids conflating epistemic and ontological reduction.

Finally, let me come back to the criteria of adequacy that I identify in Chap. 2 and show in which sense my account of explanatory reduction in biology satisfies them.

Criteria of Adequacy

An account of reduction in biology is adequate iff

(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 (**Descriptive Adequacy**),

(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 (**Balancing Specificity Against Generality**),
(3) it clarifies in which sense reductions contribute to or hinder the realization of relevant *epistemic norms*, for instance, explanatory success (**Norm-normativity**), and
(4) it is *potentially useful* to contemporary biological practice and to the broader society (**Potential Usefulness**).

The three characteristics of reductive explanation I develop are the result of an extensive analysis of paradigmatic and important examples of reductive explanations from biological practice and of discussions about reductionism that can be found in current biological research papers (criterion 1*a). 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 'What makes biological explanations reductive and distinguishes them from non-reductive ones?' is not important to them by itself.

In developing my account of explanatory reduction I try to capture much of the diversity of reductive reasoning that is present in different biological disciplines. Hence, the empirical data that I took into account came from various different biological fields. My aim is to account for this diversity while, at the same time, to reconstruct a coherent philosophical account that is as universal as possible (criterion 2). Establishing coherency involves assessing certain assumptions of biologists as incorrect or as inadequate, for instance, the view that to explain a phenomenon reductively amounts to treating the biological object or system whose behavior is to be explained as aggregative or the thesis that reductive explanations disregard all relational properties of and interactions between biological parts. The central goal of developing my account of explanatory reduction is 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.

The norms of reductive explanation are, however, not the only kind of norms that play a role in my analysis. My account of explanatory reduction is also normative in the sense that it clarifies in which sense reductions contribute to or hinder the realization of a relevant epistemic norm, namely explanatory success (criterion 3). In other words, my account also reveals the conditions under which reductive explanations are adequate, which are different from the conditions under which an explanation is reductive. That is, by identifying the three characteristics of reductive explanations I also reveal what the merits and limitations of reductive explanations are. For instance, reductive explanations are inadequate if reference to higher-level factors is necessary for explaining the behavior of the object as a whole or if the whole is only minimally decomposable or not decomposable at all, rather than being nearly decomposable.

Since my account is sensitive to contemporary biological practice it is also potentially useful to biologists (criterion 4). In clarifying the concept of a reductive explanation and specifying the strengths and limits of reductive explanations, my account has the potential to enhance debates about explanatory reductionism – not only in philosophy, but also in biological science itself.

References

Achinstein, P. (1983). *The nature of explanation*. New York: Oxford University Press.

Ahn, A. C., Tewari, M., Poon, C.-S., & Phillips, R. S. (2006a). The limits of reductionism in medicine: Could systems biology offer an alternative? *PLoS Medicine, 3*(6), 709–713.

Ahn, A. C., Tewari, M., Poon, C.-S., & Phillips, R. S. (2006b). The clinical applications of a systems approach. *PLoS Medicine, 3*(7), 1–5.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2008). *Molecular biology of the cell* (5th ed.). New York: Garland Science.

Anderson, G. (2005). Kritische oder beschreibende Wissenschaftstheorie? In B. Gesang (Ed.), *Deskriptive oder normative Wissenschaftstheorie?* (pp. 75–90). Frankfurt: Ontos.

Ankeny, R., Chang, H., Boumans, M., & Boon, M. (2011). Introduction: Philosophy of science in practice. *European Journal for Philosophy of Science, 1*, 303–307.

Ayala, F. J. (1974). Introduction. In F. J. Ayala & T. Dobzhansky (Eds.), *Studies in the philosophy of biology* (pp. vii–xvi). Berkeley: University of California Press.

Baker, L. R. (1997). Why constitution in not identity. *Journal of Philosophy, 94*(12), 599–621.

Baker, L. R. (2000). *Persons and bodies: A constitution view*. New York: Cambridge University Press.

Baker, L. R. (2007). *The metaphysics of everyday life. An essay in practical realism*. Cambridge: Cambridge University Press.

Baker, L. R. (2009). Non-reductive materialism. In B. P. McLaughlin (Ed.), *The Oxford handbook of philosophy of mind* (pp. 109–127). Oxford: Clarendon Press.

Bateson, P. (2005). The return of the whole organism. *Journal of Bioscience, 30*(1), 31–39.

Baumgartner, M., & Gebharter, A. (2015). Constitutive relevance, mutual manipulability, and fat-handedness. *The British Journal for the Philosophy of Science*. doi: 10.1093/bjps/axv003. First published online: February 20, 2015

Beatty, J. (1981). What's wrong with the received view of evolutionary theory? *PSA, 1980*(2), 397–426.

Beatty, J. (1990). Evolutionary anti–reductionism: Historical considerations. *Biology and Philosophy, 5*, 199–210.

Beatty, J. (1995). The evolutionary contingency thesis. In G. Wolters & J. Lennox (Eds.), *Concepts, theories, and rationality in the biological sciences* (pp. 45–81). Pittsburgh: University of Pittsburgh Press.

Bechtel, W. (2001). The compatibility of complex systems and reduction: A case analysis of memory research. *Minds and Machines, 11*, 483–502.

Bechtel, W. (2006). *Discovering cell mechanisms. The creation of modern cell biology*. Cambridge: Cambridge University Press.

M.I. Kaiser, *Reductive Explanation in the Biological Sciences*, History, Philosophy and Theory of the Life Sciences 16,
DOI 10.1007/978-3-319-25310-7

Bechtel, W. (2008). *Mental mechanisms. Philosophical perspectives on cognitive neuroscience*. New York: Taylor and Francis Group.

Bechtel, W. (2009). Looking down, around, and up: Mechanistic explanation in psychology. *Philosophical Psychology, 22*(5), 543–564.

Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences, 36*, 421–441.

Bechtel, W., & Abrahamsen, A. (2011). Complex biological mechanisms: Cyclic, oscillatory, and autonomous. In C. A. Hooker (Ed.), *Philosophy of complex systems. Handbook of the philosophy of science* (pp. 257–285). New York: Elsevier.

Bechtel, W., & Hamilton, A. (2007). Reduction, integration, and the unity of science: Natural, behavioral, and social sciences and the humanities. In T. A. F. Kuipers (Ed.), *General philosophy of science: Focal issues* (pp. 377–430). Amsterdam: Elsevier.

Bechtel, W., & Richardson, R. C. (2010). *Discovering complexity. Decomposition and localization as strategies in scientific research*. Cambridge: MIT Press.

Beckermann, A. (1992). Introduction – Reductive and nonreductive physicalism. In A. Beckermann, H. Flohr, & J. Kim (Eds.), *Emergence or reduction? Essays on the prospects of nonreductive physicalism* (pp. 1–21). Berlin: de Gruyter.

Bedau, M. A., & Humphreys, P. (2008). *Emergence. Contemporary readings in philosophy and science*. Cambridge: MIT Press.

Beresford, M. J. (2010). Medical reductionism: Lessons from the great philosophers. *Quarterly Journal of Medicine, 103*(9), 721–724.

Bickle, J. (1998). *Psychoneural reduction. The new wave*. Cambridge: MIT Press.

Bickle, J. (2003). *Philosophy and neuroscience. A ruthlessly reductive account*. Dordrecht: Kluwer Academic Publishers.

Bickle, J. (2006). Reducing mind to molecular pathways: Explicating the reductionism implicit in current cellular and molecular neuroscience. *Synthese, 151*, 411–434.

Bizzarri, M., Cucina, A., Conti, F., & D'Anselmi, F. (2008). Beyond the oncogene paradigm: Understanding complexity in carcinogenesis. *Acta Biotheoretica, 56*, 173–196.

Boyd, R. (1980). Materialism without reductionism: What physicalism does not entail. In N. J. Block (Ed.), *Readings in philosophy of psychology* (Vol. 1, pp. 67–106). Cambridge: Harvard University Press.

Boyle, R. (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, R. N. (1990). *Adaptation and environment*. Princeton: Princeton University Press.

Brandon, R. N. (1996). *Concepts and methods in evolutionary biology*. Cambridge: Cambridge University Press.

Brigandt, I. (2013). Explanation in biology: Reduction, pluralism, and explanatory aims. *Science and Education, 22*(1), 69–91.

Brigandt, I., & Love, A. (2008). Reductionism in biology. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy.* (Fall 2008 Edition). http://plato.stanford.edu/archives/fall2008/entries/reduction-biology/

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

Buchanan, A. V., Weiss, K. M., & Fullerton, S. M. (2006). Dissecting complex disease: The quest for the philosopher's stone? *International Journal of Epidemiology, 35*, 562–571.

Byerly, H. (2003). Reductionism: Analysis and synthesis in biological explanations. *The Quarterly Review of Biology, 78*(3), 336–342.

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

Carnap, R. (1950). *Logical foundations of probability*. Chicago: University of Chicago Press.

Carrier, M. (2007). Wege der Wissenschaftsphilosophie im 20. Jahrhundert. In A. Bartels & M. Stöckler (Eds.), *Wissenschaftstheorie. Ein Studienbuch* (pp. 15–44). Paderborn: Mentis.

Cartwright, N. (1999). *The dappled world: A study of the boundaries of science*. Cambridge: Cambridge University Press.

Casadevall, A., Fang, F. C., & Pirofski, L.-A. (2011). Microbial virulence as an emergent property: Consequences and opportunities. *PLoS Pathogens, 7*, 1–3.

Chalmers, D. J. (1996). *The conscious mind. In search of a fundamental theory*. New York: Oxford University Press.

Chalmers, D. J., & Jackson, F. (2001). Conceptual analysis and reductive explanation. *Philosophical Review, 110*(3), 315–360.

Chang, H. (2011). The philosophical grammar of scientific practice. *International Studies in the Philosophy of Science, 25*(3), 205–221.

Chong, L., & Ray, L. B. (2002). Whole-istic biology. *Science, 295*, 1661.

Coffa, J. A. (1974). Hempel's ambiguity. *Synthese, 28*(2), 141–163.

Craver, C. F. (2002a). Structures of scientific theories. In P. Machamer & M. Silberstein (Eds.), *The Blackwell guide to the philosophy of science* (pp. 55–79). Malden: Blackwell Publishers.

Craver, C. F. (2002b). Interlevel experiments and multilevel mechanisms in the neuroscience of memory. *Philosophy of Science, 69*, 83–97.

Craver, C. F. (2005). Beyond reduction: Mechanisms, multifield integration, and the unity of neuroscience. *Studies in the History and Philosophy of Biological and Biomedical Sciences, 36*, 373–395.

Craver, C. F. (2007a). *Explaining the brain. Mechanisms and the mosaic unity of neuroscience*. Oxford: Oxford University Press.

Craver, C. F. (2007b). Constitutive explanatory relevance. *Journal of Philosophical Research, 32*, 3–20.

Craver, C. F. (2013). Functions and mechanisms: A perspectivalist view. In P. Huneman (Ed.), *Functions: Selection and mechanisms* (Synthese library, pp. 133–158). Dordrecht: Springer.

Craver, C. F. (2014). The ontic conception of scientific explanation. In A. Hüttemann, M. I. Kaiser, & O. Scholz (Eds.), *Explanation in the special sciences – The case of biology and history* (Synthese library, pp. 27–52). Dordrecht: Springer.

Craver, C. F., & Bechtel, W. (2007). Top-down causation without top-down causes. *Biology and Philosophy, 22*, 547–563.

Craver, C. F., & Darden, L. (2001). Discovering mechanisms in neurobiology: The case of spatial memory. In P. Machamer, R. Grush, & P. McLaughlin (Eds.), *Theory and method in neuroscience* (pp. 112–137). Pittsburgh: University of Pittsburgh Press.

Craver, C. F., & Darden, L. (2005). Introduction. *Studies in History and Philosophy of Biological and Biomedical Sciences, 36*, 233–244.

Craver, C. F., & Darden, L. (2013). *In search of mechanisms. Discoveries across the life sciences*. Chicago: University of Chicago Press.

Crick, F. H. C. (1966). *Of molecules and men*. Seattle: University of Washington Press.

Crick, F. H. C. (1988). *What mad pursuit. A personal view of scientific discovery*. New York: Francis Books.

Cummins, R. C. (1975). Functional analysis. *The Journal of Philosophy, 72*(20), 741–765.

Cummins, R. C. (1983). *The nature of psychological explanation*. Cambridge: MIT Press.

Darden, L. (1991). *Theory change in science. Strategies from mendelian genetics*. New York: Oxford University Press.

Darden, L. (2005). Relations among fields: Mendelian, cytological and molecular mechanisms. *Studies in History and Philosophy of Biological and Biomedical Sciences, 36*, 357–371.

Darden, L. (2008). Thinking again about biological mechanisms. *Philosophy of Science, 75*, 958–969.

Darden, L., & Maull, N. (1977). Interfield theories. *Philosophy of Science, 44*, 43–64.

Darden, L., & Tabery, J. (2010). Molecular biology. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2010 Edition). http://plato.stanford.edu/archives/fall2010/entries/molecular-biology/

Davidson, D. (1970). Mental events. In L. Foster & J. W. Swanson (Eds.), *Experience and theory* (pp. 79–101). London: Duckworth.

Dawkins, R. (1976). *The selfish gene*. New York: Oxford University Press.

Dawkins, R. (1982). *The extended phenotype: The gene as the unit of selection*. Oxford: Freeman.

De Vreese, L., Weber, E., & Van Bouwel, J. (2010). Explanatory pluralism in the medical science: Theory and practice. *Theoretical Medicine and Bioethics, 31*, 371–390.

Delehanty, M. (2005). Emergent properties and the context objection to reduction. *Biology and Philosophy, 20*, 715–734.

Dennett, D. C. (1995). *Darwin's dangerous idea*. New York: Simon and Schuster.

Descartes, R. (1991 [1644]). *Principles of philosophy* (Translated, with explanatory notes, V. R. Miller & R. P. Miller), Dordrecht: Kluwer.

Dizadji-Bahmani, F., Frigg, R., & Hartmann, S. (2010). Who is afraid of Nagelian reduction? *Erkenntnis, 73*, 393–412.

Dobson, C. M. (2003). Protein folding and misfolding. *Nature, 426*, 884–890.

Dobzhansky, T. (1973). Nothing in biology makes sense except in the light of evolution. *American Biology Teacher, 35*, 125–129.

Dowe, P. (1992). Process causality and asymmetry. *Erkenntnis, 37*(2), 179–196.

Dowe, P. (2000). *Physical causation*. Cambridge: Cambridge University Press.

Dupré, J. (1993). *The disorder of things. Metaphysical foundations of the disunity of science*. Cambridge: Harvard University Press.

Dupré, J. (2009). It is not possible to reduce biological explanations to explanations in chemistry and/ or physics. In F. J. Ayala & R. Arp (Eds.), *Contemporary debates in philosophy of biology* (pp. 32–47). Chichester: Blackwell.

Dupré, J. (2012). *Processes of life. Essays in the philosophy of biology*. Oxford: Oxford University Press.

Einstein, A. (1933). *On the method of theoretical physics*. The Herbert Spencer lecture, delivered at Oxford, 10 June 1933, Oxford: Clarendon Press.

Eronen, M. I. (2015). Levels of organization: A deflationary account. *Biology and Philosophy, 30*(1), 39–58.

Falkenburg, B. (2005). Der Wert wertfreier Wissenschaft. In B. Gesang (Ed.), *Deskriptive oder normative Wissenschaftstheorie?* (pp. 91–122). Frankfurt: Ontos.

Fang, F. C., & Casadevall, A. (2011). Reductionistic and holistic science. *Infection and Immunity, 79*, 1401–1404.

Fazekas, P., & Kertész, G. (2011). Causation at different levels: Tracking the commitments of mechanistic explanations. *Biology and Philosophy, 26*, 365–383.

Feigl, H. (1958). The 'Mental' and the 'Physical'. In H. Feigl, M. Scriven, & G. Maxwell (Eds.), *Concepts, theories and the mind-body problem* (Minnesota studies in the philosophy of science, Vol. 2). Minneapolis: University of Minnesota Press.

Feyerabend, P. (1962). Explanation, reduction and empiricism. In H. Feigl & G. Maxwell (Eds.), *Scientific explanation, space, and time* (pp. 28–97). Minneapolis: University of Minnesota Press.

Fincham, J. R. S. (2000). Reductionism should be clarified, not dismissed. *Nature, 406*, 343.

Fodor, J. A. (1974). Special sciences (or: The disunity of science as a working hypothesis). *Synthese, 28*, 97–115.

Forge, J. (1998). Explanation and mechanism: Reflections on the ontic conception of explanation. In D. Anapolitanos, A. Baltas, & S. Tsinorema (Eds.), *Philosophy and the many faces of science* (pp. 76–92). Lanham: Rowman & Littelfield Publishers.

Friedman, M. (1974). Explanation and scientific understanding. *Journal of Philosophy, 71*(1), 5–19.

Frost-Arnold, G. (2004). How to be an anti-reductionist about developmental biology: Response to Laubichler and Wagner. *Biology and Philosophy, 19*, 75–91.

Frydman, J. (2001). Folding of newly translated proteins *in vitro*: The role of molecular chaperones. *Annual Review of Biochemistry, 70*, 603–647.

Gallagher, R., & Appenzeller, T. (1999). Beyond reductionism. *Science, 284*, 79.

Gesang, B. (2005). Normative Wissenschaftstheorie – Ein längst verstorbener Patient? In B. Gesang (Ed.), *Deskriptive oder normative Wissenschaftstheorie?* (pp. 11–30). Frankfurt: Ontos.

Giere, R. N. (1988). *Explaining science. A cognitive approach*. Chicago: University of Chicago Press.

Giere, R. N. (1999). *Science without laws*. Chicago: University of Chicago Press.

Gilbert, S. F. (2006). *Developmental biology* (8th ed.). Sunderland: Sinauer Associates.

Gillett, C. (2007). Understanding the reductionism: The metaphysics of science and compositional reduction. *Journal of Philosophy, 104*(4), 193–216.

Gillett, C. (2013). Constitution, and multiple constitution, in the sciences: Using the neuron to construct a starting framework. *Minds and Machines, 23*(3), 309–337.

Glennan, S. S. (2002). Rethinking mechanistic explanation. *Philosophy of Science, 69*, 342–353.

Glennan, S. S. (2008). Mechanism. In S. Psillos & M. Curd (Eds.), *The Routledge companion to philosophy of science* (pp. 376–384). London: Routledge.

Glennan, S. S. (2010). Ephemeral mechanisms and historical explanation. *Erkenntnis, 72*(2), 251–266.

Godfrey-Smith, P. (2008). Reduction in real life. In J. Hohwy & J. Kallestrup (Eds.), *Being reduced. New essays on reduction, explanation and causation* (pp. 52–74). Oxford: Oxford University Press.

Greenspan, R. J. (2001). The flexible genome. *Nature Reviews, 2*, 383–387.

Griffiths, P. E. (2007). Philosophy of biology. In S. Sarkar & J. Pfeifer (Eds.), *The philosophy of science. An encyclopedia* (pp. 68–75). New York: Routledge.

Grizzi, F., & Chiriva-Internati, M. (2005). The complexity of anatomical systems. *Theoretical Biology and Medical Modelling, 2*, 26–34.

Grizzi, F., & Chiriva-Internati, M. (2006). Cancer: Looking for simplicity and finding complexity. *Cancer Cell International, 6*, 4–10.

Grizzi, F., Di Ieva, A., Russo, C., Frezza, E. E., Cobos, E., Muzzio, P. C., & Chiriva-Internati, M. (2006). Cancer initiation and progression: An unsimplifiable complexity. *Theoretical Biology and Medical Modelling, 3*, 37–42.

Grush, R. (2003). In defense of some 'Cartesian' assumption concerning the brain and its operation. *Biology and Philosophy, 18*, 53–93.

Guttman, B. S. (1976). Is 'Levels of Organization' a useful biological concept? *BioScience, 26*(2), 112–113.

Hahn, W. C., & Weinberg, R. A. (2002). Modelling the molecular circuitry of cancer. *Nature Reviews, 2*, 331–341.

Hartl, F. U. (2011). Chaperone-assisted protein folding: The path to discovery from a personal perspective". *Nature Medicine, 17*, 1206–1210.

Hartl, F. U., & Hayer-Hartl, M. (2002). Molecular chaperones in the cytosol: From nascent chain to folded protein. *Science, 295*, 1852–1858.

Haugeland, J. (1998). *Having thought. Essays in the metaphysics of mind*. Cambridge: Harvard University Press.

Hempel, C. G. (1965). *Aspects of scientific explanation. And other essays in the philosophy of science*. New York: Free.

Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science, 15*(2), 135–175.

Hitchcock, C. (1995). Discussion: Salmon on explanatory relevance. *Philosophy of Science, 62*, 304–20.

Hitchcock, C., & Woodward, J. (2003). Explanatory generalizations, part II: Plumbing explanatory depth. *Noûs, 37*(2), 181–199.

Hooker, C. (1981). Towards a general theory of reduction. Part I: Historical and scientific setting. Part II: Identity in reduction. Part III: Cross-categorical reduction. *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. *Nature, 421*, 172–177.

Hoyningen-Huene, P. (2007). Reduktion und Emergenz. In A. Bartels & M. Stöckler (Eds.), *Wissenschaftstheorie. Ein Studienbuch* (pp. 177–197). Paderborn: Mentis.

Hull, D. (1972). Reductionism in genetics – Biology or philosophy? *Philosophy of Science, 39*, 491–499.

Hull, D. (1974). *The philosophy of biological science*. Englewood Cliffs: Prentice-Hall.

Hull, D. (1976). Informal aspects of theory reduction. *PSA, 1974*, 653–670.

Hull, D. L., & van Regenmortel, M. H. V. (2002). Introduction. In M. H. V. van Regenmortel & D. L. Hull (Eds.), *Reductionism in the biomedical sciences* (pp. 1–13). London: Wiley.

Huneman, P. (2010). Topological explanations and robustness in biological sciences. *Synthese, 177*, 213–245.

Huneman, P. (2014a). Individuality as a theoretical scheme. I. Formal and material concepts of individuality. *Biological Theory, 9*(4), 361–373.

Huneman, P. (2014b). Individuality as a theoretical scheme. II. About the weak individuality of organisms and ecosystems. *Biological Theory, 9*(4), 374–381.

Hunter, P. (2003). Putting Humpty Dumpty back together again. *Scientist, 17*(4), 20–22.

Hüttemann, A. (2004). *What's wrong with microphysicalism?* New York: Routledge.

Hüttemann, A., & Love, A. C. (2011). Aspects of reductive explanation in biological science: Intrinsicality, fundamentality, and temporality. *British Journal for Philosophy of Science, 62*(3), 519–549.

Huxley, H. E., & Hanson, J. (1954). Changes in the cross-striations of muscle during contraction and stretch and their structural interpretation. *Nature, 1973*, 973–976.

Huxley, A. F., & Niedergerke, R. (1954). Structural changes in muscle during contraction; interference microscopy of living muscle fibres. *Nature, 1973*, 971–973.

Illari, P. (2013). Mechanistic explanation: Integrating the ontic and epistemic. *Erkenntnis, 78*(2), 237–255.

Illari, P., & Williamson, J. (2012). What is a mechanism? Thinking about mechanisms across the sciences. *European Journal of Philosophy of Science, 2*, 119–135.

Jackson, F. (1998). *From metaphysics to ethics. A defense of conceptual analysis*. Oxford: Oxford University Press.

Janich, P. (2005). Wissenschaftsphilosophie als kritische Reflexion auf eine historische Praxis. In B. Gesang (Ed.), *Deskriptive oder normative Wissenschaftstheorie?* (pp. 145–166). Frankfurt: Ontos.

Jansen, L., & Schulz, S. (2014). Crisp islands in vague seas: Cases of determinate parthood relations in biological objects. In C. Calosi & P. Graziani (Eds.), *Mereology in the sciences. Parts and wholes in contemporary scientific contexts* (pp. 163–188). Cham: Springer.

Johnston, M. (1992). Constitution is not identity. *Mind, 101*(401), 89–106.

Joyner, M. J., & Pedersen, B. K. (2011). Ten questions about systems biology. *The Journal of Physiology, 589*(5), 1017–1030.

Kaiser, M. I. (2011). Limits of reductionism in the life sciences. *History and Philosophy of the Life Sciences, 33*, 453–476.

Kaiser, M. I. (2012). Why it is time to move beyond nagelian reduction. In D. Dieks, W. J. Gonzalez, S. Hartmann, M. Stöltzner, & M. Weber (Eds.), *Probabilities, laws, and structures* (The philosophy of science in a European perspective, Vol. 3, pp. 255–272). Heidelberg: Springer.

Kaiser, M. I. (2015). Philosophy of microbiology – Maureen A. O'Malley. *International Studies in the Philosophy of Science, 29*(2), doi: 10.1080/02698595.2015.1119427.

Kaiser, M. I. (forthcoming a). Individuating part-whole relations in the biological world. In O. Bueno, R.-L. Chen, & M. B. Fagan (Eds.), *Individuation across experimental and theoretical sciences*, Oxford University Press.

Kaiser, M. I. (forthcoming b). Biological parts. In: H. Burkhardt, J. Seibt, G. Imaguire (Eds.), *Handbook of mereology*, Philosophia.

Kaiser, M. I. (manuscript a). *Levels as determined by parts, wholes, and kinds*

Kaiser, M. I. (manuscript b) *ENCODE and the question of Parthood*.

Kaiser, M. I., & Craver, C. F. (2013). Mechanisms and laws: Clarifying the debate. In H.-K. Chao, S.-T. Chen, & R. L. Millstein (Eds.), *Mechanism and causality in biology and economics* (pp. 125–145). Dordrecht: Springer.

Kaiser, M. I., & Krickel, B. (forthcoming). The metaphysics of constitutive mechanistic phenomena. *British Journal for the Philosophy of Science*.

Kauffmann, S. A. (1970). Articulation of parts explanation in biology and the rational search for them. *PSA, 1970*, 257–272.

Keil, G., & Schnädelbach, H. (Eds.). (2000). *Naturalismus. Philosophische Beiträge*. Frankfurt am Main: Suhrkamp.

Keller, F. E. (2005). The century beyond the gene. *Journal of Bioscience, 30*(1), 3–10.

Kellert, S. H., Longino, H. E., & Waters, C. K. (2006). Introduction: The pluralist stance. In S. H. Kellert, H. E. Longino, & C. K. Waters (Eds.), *Scientific pluralism. Minnesota studies in the philosophy of science* (pp. vii–xxix). Minneapolis: University of Minnesota Press.

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. *Genes & Development, 11*, 701–713.

Kim, J. (1989). The myth of non-reductive materialism. *Proceedings and Addresses of the American Philosophical Association, 63*(3), 31–47.

Kim, J. (1993). *Mind and supervenience*. Cambridge: Cambridge University Press.

Kim, J. (1999). Making sense of emergence. *Philosophical Studies, 95*, 3–36.

Kim, J. (2005). *Physicalism, or something near enough*. Princeton: Princeton University Press.

Kim, J. (2006). *Philosophy of mind* (2nd ed.). Oxford: Westview.

Kim, J. (2008). Reduction and reductive explanation: Is one possible without the other? In J. Hohwy & J. Kallestrup (Eds.), *Being reduced. New essays on reduction, explanation, and causation* (pp. 91–114). Oxford: Oxford University Press.

Kincaid, H. (1990). Molecular biology and the unity of science. *Philosophy of Science, 57*(4), 575–593.

Kitano, H. (2002). Systems biology: A brief overview. *Science, 295*, 1662–1664.

Kitcher, P. (1981). Explanatory unification. *Philosophy of Science, 84*, 507–531.

Kitcher, P. (1984). 1953 an all that: A tale of two sciences. *Philosophical Review, 93*, 335–373.

Kitcher, P. (1989). Explanatory unification and the causal structure of the world. In P. Kitcher & W. C. Salmon (Eds.), *Scientific explanation* (pp. 410–506). Minneapolis: University of Minnesota Press.

Kitcher, P. (1993). Function and design. *Midwest Studies in Philosophy, 18*(1), 379–397.

Kitcher, P. (1999a). The hegemony of molecular biology. *Biology and Philosophy, 14*, 195–210.

Kitcher, P. (1999b). Unification as a regulative ideal. *Perspectives on Science, 7*(3), 337–348.

Kitcher, P. (2011). Philosophy inside out. *Metaphilosophy, 42*(3), 248–260.

Kitcher, P., & Salmon, W. C. (1987). Van Fraassen on explanation. *The Journal of Philosophy, 84*(6), 315–330.

Klein, C. (2009). Reduction without reductionism: A defense of Nagel on connectability. *The Philosophical Quarterly, 59*, 39–53.

Korfiatis, K. J., & Stamou, G. P. (1999). Habitat templets and the changing worldview of ecology. *Biology and Philosophy, 14*, 375–393.

Kornblith, H. (2007). Naturalism and intuitions. *Grazer Philosophische Studien, 74*(1), 27–49.

Krohs, U. (2004). *Eine Theorie Biologischer Theorien. Status und Gehalt von Funktionsaussagen und informationstheoretischen Modellen*. Berlin: Springer.

Ladyman, & J., Ross (2007). *Every thing must go. Metaphysics naturalized*. Oxford: Oxford University Press.

Lange, M. (2000). *Natural laws in scientific practice*. Oxford: Oxford University Press.

Laubichler, M. D., & Wagner, G. P. (2001). How molecular is molecular developmental biology? A reply to Alex Rosenberg's 'Reductionism Redux: Computing the Embryo'. *Biology and Philosophy, 16*, 53–68.

Leuridan, B. (2012). Three problems for the mutual manipulability account of constitutive relevance in mechanisms. *British Journal of Philosophy of Science, 63*(2), 1–29.

Levenstein, S. (2009). Against reductionism. *BMJ, 339*, 709.

Levins, R. (1970). Complex systems. In C. H. Waddington (Ed.), *Towards a theoretical biology* (pp. 73–88). Edinburgh: University Press.

Levins, R., & Lewontin, R. C. (1980). Dialectics and reductionism in ecology. *Synthese, 43*, 47–78.

Lewis, D. (1983). New work for a theory of universals. *Australasian Journal of Philosophy, 61*(4), 343–377.

Lewis, D. (1986). Events. In D. Lewis (Ed.), *Philosophical papers* (Vol. II, pp. 241–269). Oxford: Oxford University Press.

Lewis, D. (1994). Reduction in mind. In S. Guttenplan (Ed.), *A companion to philosophy of mind* (pp. 412–431). Oxford: Blackwell.

Lidicker, W. Z. (1988). The synergistic effects of reductionist and holistic approaches in animal ecology. *OIKOS, 53*, 278–281.

Lloyd, E. (1988). *The structure and confirmation of evolutionary theory*. New York: Greenwood Press.

Loehle, C. (1988). Philosophical tools: Potential contributions to ecology. *OIKOS, 51*(1), 97–119.

Love, A. C. (2008a). From philosophy to science (to natural philosophy): Evolutionary developmental perspectives. *The Quarterly Review of Biology, 83*(1), 65–76.

Love, A. C. (2008b). Review symposium: More worry and less love. *Metascience, 17*, 1–26.

Love, A. C. (2012a). Formal and material theories in philosophy of science: A methodological interpretation". In H. K. de Regt, S. Hartmann, & S. Okasha (Eds.), *EPSA philosophy of science: Amsterdam 2009* (pp. 175–185). Dordrecht: Springer.

Love, A. C. (2012b). Hierarchy, causation and explanation: Ubiquity, locality and pluralism. *Interface Focus, 2*, 115–125.

Love, A. C., & Hüttemann, A. (2011). Comparing part-whole reductive explanations in biology and physics. In D. Dieks, W. J. Gonzalez, S. Hartmann, T. Uebel, & M. Weber (Eds.), *Explanation, prediction, and confirmation. New trends and old ones reconsidered* (pp. 183–202). Berlin: Springer.

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science, 67*, 1–25.

Macilwain, C. (2009). Watching science at work. *Nature, 462*, 840–842.

Mackie, J. L. (1974). *The cement of the universe. A study of causation*. Oxford: Clarendon Press.

Malaterre, C. (2007). Organicism and reductionism in cancer research: Towards a systemic approach. *International Studies in the Philosophy of Science, 21*(1), 57–73.

Marcum, J. A. (2005). Metaphysical presuppositions and scientific practices: Reductionism and organicism in cancer research. *International Studies in the Philosophy of Science, 19*(1), 31–45.

Mayr, E. (1982). *The growth of biological thought. Diversity, evolution, and inheritance*. Cambridge: Belknap Press.

Mayr, E. (1988). The limits of reductionism. *Nature, 331*, 475–476.

Mayr, E. (2004). *What makes biology unique? Considerations on the autonomy of a scientific discipline*. Cambridge: Cambridge University Press.

Mazzocchi, F. (2008). Complexity in biology. Exceeding the limits of reductionism and determinism using complexity theory. *EMBO Reports, 9*, 10–14.

McClelland, E. E., Bernhardt, P., & Casadevall, A. (2005). Coping with multiple virulence factors: Which is most important? *PLoS Pathogens, 1*, 287–288.

McLaughlin, P. (2001). *What functions explain: Functional explanation and self-reproducing systems*. Cambridge: Cambridge University Press.

McLaughlin, B., & Bennett, K. (2011). Supervenience. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2011 Edition). http://plato.stanford.edu/archives/win2011/entries/supervenience/

McShea, D. W. (2000). Functional complexity in organisms: Parts as proxies. *Biology and Philosophy, 15*, 641–668.

McShea, D. W., & Venit, E. (2001). What is a part? In G. P. Wagner (Ed.), *The character concept in evolutionary biology* (pp. 259–284). San Diego: Academic.

Mellor, D. H. (2008). Micro-composition. *Royal Institute of Philosophy Supplements, 83*(62), 65–80.

Mikkelson, G. M. (2004). Biological diversity, ecological stability, and downward causation. In M. Oksanen & J. Pietarinen (Eds.), *Philosophy and biodiversity* (pp. 119–129). Cambridge: Cambridge University Press.

Mikkelson, G. M. (forthcoming). Part-whole relationships and the unity of ecology. In R. A. Skipper Jr., C. Allen, R. Ankeny, C. F. Craver, L. Darden, G. M. Mikkelson, & R. C. Richardson (Eds.), *Philosophy across the life sciences*. MIT Press.

Mitchell, S. D. (1997). Pragmatic laws. *Philosophy of Science, 64*, 468–479.

Mitchell, S. D. (2003). *Biological complexity and integrative pluralism*. Cambridge: Cambridge University Press.

Mitchell, S. D. (2009). *Unsimple truths. Science, complexity, and policy*. Chicago: University of Chicago Press.

Mitchell, S. D., & Dietrich, M. (2006). Integration without unification: An argument for pluralism in the biological sciences. *American Naturalist, 168*, 73–79.

Morgan, M., & Morrison, M. (1999). *Models as mediators. Perspectives on natural and social science*. Cambridge: Cambridge University Press.

Mühlhölzer, F. (2005). Naturalismus und Lebenswelt – Plädoyer für eine rein deskriptive Wissenschaftstheorie. In B. Gesang (Ed.), *Deskriptive oder normative Wissenschaftstheorie?* (pp. 49–73). Frankfurt: Ontos.

Nagel, E. (1952). Wholes, sums, and organic unities. *Philosophical Studies, 3*(2), 17–32.

Nagel, E. (1961). *The structure of science. Problems in the logic of scientific explanation*. London: Routledge.

Nickles, T. (1973). Two concepts of intertheoretic reduction. *The Journal of Philosophy, 70*(7), 181–201.

Noble, D. (2002). Modeling the heart – From genes to cells to the whole organ. *Sciene, 295*, 1678–1682.

Norton, J. (2003). A material theory of induction. *Philosophy of Science, 70*(4), 647–670.

O'Malley, M. A. (2014). *Philosophy of microbiology*. Cambridge: Cambridge University Press.

O'Malley, M. A., & Dupré, J. (2005). Fundamental issues in systems biology. *BioEssays, 27*, 1270–1276.

O'Malley, M. A., & Dupré, J. (2007). Size doesn't matter: Towards a more inclusive philosophy of biology. *Biology and Philosophy, 22*, 155–191.

Oppenheim, P., & Putnam, H. (1958). Unity of science as a working hypothesis. In H. Feigl, M. Scriven, & G. Maxwell (Eds.), *Concepts, theories and the mind-body problem* (Minnesota studies in the philosophy of science, Vol. II, pp. 3–36). Minneapolis: University of Minnesota Press.

Papineau, D. (2009). Naturalism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2009 Edition). http://plato.stanford.edu/archives/spr2009/entries/naturalism/

Pereboom, D. (2002). Robust nonreductive materialism. *Journal of Philosophy, 99*, 499–531.

Pereboom, D., & Kornblith, H. (1990). The metaphysics of irreducibility. *Philosophical Studies, 63*, 125–145.

Perini, L. (2005). Explanation in two dimensions: Diagrams and biological explanation. *Biology and Philosophy, 20*, 257–269.

Perini, L. (2013). Diagrams in biology. *Knowledge Engineering Review, 28*(3), 273–286.

Place, U. T. (1956). Is consciousness a brain process? *British Journal of Psychology, 47*(1), 44–50.

Plantinga, A. (1996). Methodological naturalism? In J. Van der Meer (Ed.), *Facets of faith and science*. Lanham: University Press of America.

Popper, K. R. (1974). Scientific reduction and the essential incompleteness of all science. In F. J. Ayala & T. G. Dobzhansky (Eds.), *Studies in the philosophy of biology. Reduction and related problems* (pp. 259–282). Berkeley: University of California Press.

Potochnik, A. (2009). Levels of explanation reconceived. *Philosophy of Science, 77*, 59–72.

Potochnik, A., & McGill, B. (2012). The limitations of hierarchical organization. *Philosophy of Science, 79*(1), 120–140.

Powell, K. (2004). All systems go. *The Journal of Cell Biology, 165*, 299–303.

Putnam, H. (1975). *Mind, language, and reality* (Philosophical papers, Vol. 2). New York: Cambridge University Press.

Railton, P. (1981). Probability, explanation, and information. *Synthese, 48*, 233–256.

Reece, J. B., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., & Jackson, R. B. (2011). *Campbell biology* (9th ed.). Boston: Pearson.

Reichenbach, H. (1938). On probability and induction. *Philosophy of Science, 5*(1), 21–45.

Richfield, D. (2014). Medical gallery of David Richfield. *Wikiversity Journal of Medicine*, *1*(2). doi:10.15347/wjm/2014.009. ISSN 2001–8762.

Rosenberg, A. (1985). *The structure of biological science*. Cambridge: Cambridge University Press.

Rosenberg, A. (1994). *Instrumental biology or the disunity of science*. Chicago: University of Chicago Press.

Rosenberg, A. (1997). Reductionism redux: Computing the embryo. *Biology and Philosophy, 12*, 445–470.

Rosenberg, A. (2001). How is biological explanation possible? *British Journal for Philosophy of Science, 52*, 735–760.

Rosenberg, A. (2006). *Darwinian reductionism. Or, how to stop worrying and love molecular biology*. Cambridge: University of Chicago Press.

Rosenberg, A., & Kaplan, D. M. (2005). How to reconcile physicalism and antireductionism about biology. *Philosophy of Science, 72*, 43–68.

Rosenberg, A., & McShea, D. W. (2008). *Philosophy of biology. A contemporary introduction*. New York: Routledge.

Roukos, D. H. (2011). Networks medicine: From reductionism to evidence of complex dynamic biomolecular interactions. *Pharmacogenomics, 12*(5), 695–698.

Ruse, M. (1976). Reduction in genetics. *PSA, 1974*, 633–651.

Salmon, W. C. (1971). Statistical explanation. In W. C. Salmon (Ed.), *Statistical explanation and statistical relevance* (pp. 29–87). Pittsburgh: University of Pittsburgh Press.

Salmon, W. C. (1977). A third dogma of empiricism. In R. Butts & J. Hintikka (Eds.), *Basic problems in methodology and linguistics* (pp. 149–166). Dordrecht: Reidel.

Salmon, W. C. (1984a). *Scientific explanation and the causal structure of the world*. Princeton: Princeton University Press.

Salmon, W. C. (1984b). Scientific explanation: Three basic conceptions. *PSA, 1984*(2), 293–305.

Salmon, W. C. (1989). Four decades of scientific explanation. In P. Kitcher & W. C. Salmon (Eds.), *Scientific explanation* (pp. 3–219). Minneapolis: University of Minnesota Press.

Salmon, W. C. (1992). Scientific explanation. In M. H. Salmon, J. Earman, C. Glymour, J. G. Lennox, P. Machamer, J. E. McGuire, J. D. Norton, W. C. Salmon, & K. F. Schaffner (Eds.), *Introduction to the philosophy of science* (pp. 7–41). Indianapolis: Hackett Publishing.

Salmon, W. C. (1993). Scientific explanation and the causal structure of the world. In D. Hillel-Ruben (Ed.), *Explanation. Oxford readings in philosophy* (pp. 78–112). Oxford: Oxford University Press.

Salmon, W. C. (1994). Causality without counterfactuals. *Philosophy of Science, 61*, 297–312.

Salmon, W. C. (1997). Causality and explanation: A reply to two critiques. *Philosophy of Science, 64*, 461–477.

Salmon, W. C. (1998). *Causality and explanation*. Oxford: Oxford University Press.

Sankey, H. (2013). On the evolution of criteria of theory choice. *Metascience, 22*(1), 169–172.

Sarkar, S. (1992). Models of reduction and categories of reductionism. *Synthese, 91*, 167–194.

Sarkar, S. (1998). *Genetics and reductionism*. Cambridge: Cambridge University Press.

Sarkar, S. (2005). *Molecular models of life. Philosophical papers on molecular biology*. Cambridge: MIT Press.

Sarkar, S. (2008). Reduction. In S. Psillos & M. Curd (Eds.), *The Routledge companion to philosophy of science* (pp. 425–434). London: Routledge.

Sarkar, S. (2009). Ecology. In E. N. Zalta (Ed.), *Stanford encyclopedia of philosophy* (Spring 2009 Edition). http://plato.stanford.edu/archives/spr2009/entries/ecology/

Schaffner, K. F. (1967). Approaches to reduction. *Philosophy of Science, 34*, 137–147.

Schaffner, K. F. (1969). The Watson-Crick model and reductionism. *British Journal for the Philosophy of Science, 20*, 325–348.

Schaffner, K. F. (1974a). The peripherality of reductionism in the development of molecular biology. *Journal of the History of Biology, 7*(1), 111–139.

Schaffner, K. F. (1974b). Reductionism in biology: Prospects and problems. *PSA, 1974*, 613–632.

Schaffner, K. F. (1993). *Discovery and explanation in biology and medicine*. Chicago: University of Chicago Press.

Schaffner, K. F. (2006). Reduction: The Cheshire cat problem and a return to the roots. *Synthese, 151*, 377–402.

Schindler, S. (2013). The Kuhnian mode of HPS. *Synthese, 190*, 4137–4154.

Schoener, T. W. (1986). Mechanistic approaches to community ecology: A new reductionism? *American Zoologist, 26*, 81–106.

Schurz, G. (2005). Rationale Rekonstruktion: die Methode der Wissenschaftstheorie. In B. Gesang (Ed.), *Deskriptive oder normative Wissenschaftstheorie?* (pp. 123–144). Frankfurt: Ontos.

Service, R. F. (1999). Exploring the systems life. *Science, 284*, 80–83.

Shimamura, H., Terada, Y., Okado, T., Tanaka, H., Inoshita, S., & Sasaki, S. (2003). The PI3-kinase-Akt pathway promotes mesangial cell survival and inhibits apoptosis in vitro via NF-kappa B and bad. *Journal of American Society of Nephrology, 14*(6), 1427–1434.

Sider, T. (2001). *Four dimensionalism: An ontology of persistence and time*. Oxford: Oxford University Press.

Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society, 106*(6), 467–482.

Simon, H. A. (1973). The organization of complex systems. In H. H. Pattee (Ed.), *Hierarchy theory: The challenge of complex systems*. New York: Braziller.

Simons, P. (1987). *Parts. A study in ontology*. Oxford: Clarendon.

Skipper, R. A., & Millstein, R. L. (2005). Thinking about evolutionary mechanisms: Natural selection. *Studies in the History and Philosophy of Biological and Biomedical Sciences, 36*, 327–347.

Sloep, P., & Van der Steen, W. (1987). The nature of evolutionary theory: The semantic challenge. *Biology and Philosophy, 2*, 1–15.

Smart, J. J. C. (1959). Sensations and brain processes. *Philosophical Review, 68*, 141–56.

Smith, B., & Varzi, A. C. (2000). Fiat and bona-fide boundaries. *Philosophy and Phenomenological Research, 60*(2), 401–420.

Sober, E. (1999). The multiple realizability argument against reductionism. *Philosophy of Science, 66*, 542–564.

Sober, E. (2008). *Evidence and evolution. The logic behind the science*. Cambridge: Cambridge University Press.

Soler, L., Zwart, S., Lynch, M., & Isreal-Jost, V. (Eds.). (2014). *Science after the practice turn in the philosophy, history, and social studies of science*. New York: Routledge.

Sonnenschein, C., & Soto, A. M. (2000). Somatic mutation theory of carcinogenesis: Why it should be dropped and replaced. *Molecular Carcinogenesis, 29*, 205–211.

Sorger, P. K. (2005). A reductionist's systems biology. *Current Opinion in Cell Biology, 17*, 9–11.

Soto, A. M., & Sonnenschein, C. (2005). Emergentism as a default: Cancer as a problem of tissue organization. *Journal of Bioscience, 30*(1), 103–118.

Soto, A. M., & Sonnenschein, C. (2006). Emergentism by default: A view from the bench. *Synthese, 151*, 361–376.

Soto, A. M., & Sonnenschein, C. (2010). Environmental causes of cancer: Endocrine disruptors as carcinogenes. *Nature Reviews, 6*, 363–370.

Soto, A. M., Rubin, B. S., & Sonnenschein, C. (2009). Interpreting endocrine disruption from an integrative biology perspective. *Molecular and Cellular Endocrinology, 304*, 3–7.

Spencer, M. (1997). The effects of habitat size and energy on food web structure: An individual-based cellular automata model. *Ecological Modelling, 94*, 299–316.

Steel, D. P. (2004). Can a reductionist be a pluralist? *Biology and Philosophy, 19*, 55–73.

Stoljar, D. (2009). Physicalism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2009 Edition). http://plato.stanford.edu/archives/fall2009/entries/physicalism/

Stotz, K., & Griffiths, P. 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*. http://www.pitt.edu/~kstotz/genes/genes.html

Stotz, K., & Griffiths, P. E. (2005). Genes: Philosophical analyses put to the test. *History and Philosophy of the Life Sciences, 26*, 5–28.

Strange, K. (2005). The end of 'Naive Reductionism': Rise of systems biology or renaissance of physiology? *American Journal of Physiology - Cell Physiology, 288*, 968–974.

Strevens, M. (2006). Scientific explanation. In D. M. Borchert (Ed.), *Encyclopedia of philosophy* (2nd ed.). Detroit: Macmillan.

Strevens, M. (2008). *Depth. An account of scientific explanations*. Cambridge: Harvard University Press.

Suppe, F. (1977). *The structure of scientific theories* (2nd ed.). Urbana: University of Illinois Press.

Suppe, F. (1989). *The semantic conception of theories and scientific realism*. Chicago: University of Illinois Press.

Suppe, F. (2000). Understanding scientific theories: An assessment of developments, 1969–1998. *Philosophy of Science, 67*, 102–115.

Tabery, J. G. (2004). Synthesizing activities and interactions in the concept of mechanism. *Philosophy of Science, 71*, 1–15.

Thompson, P. (1989). *The structure of biological theories*. Albany: State University of New York Press.

Torres, P. J. (2008). A modified conception of mechanism. *Erkenntnis, 71*(2), 233–251.

van Fraassen, B. C. (1977). The pragmatics of explanation. *American Philosophical Quarterly, 14*(2), 143–150.

van Fraassen, B. C. (1980). *The scientific image*. Oxford: Oxford University Press.

van Inwagen, P. (1990). *Material beings*. New York: Cornell University Press.

van Regenmortel, M. H. V. (1998). From absolute to exquisite specificity. Reflection on the fuzzy nature of species, specificity and antigenic sites. *Journal of Immunological Methods, 216*, 37–48.

van Regenmortel, M. H. V. (2004a). Reductionism and complexity in molecular biology. *EMBO Reports, 5*, 1016–1020.

van Regenmortel, M. H. V. (2004b). Biological complexity emerges from the ashes of genetic reductionism. *Journal of Molecular Recognition, 17*(3), 145–148.

van Regenmortel, M. H. V., & Hull, D. L. (Eds.). (2002). *Reductionism in the biomedical sciences*. London: Wiley.

Venter, J. C., et al. (2001). The sequence of the human genome. *Science, 291*, 1304–1348.

Waters, C. K. (1990). Why the antireductionist consensus won't survive the case of classical Mendelian genetics. *PSA, 1990*(1), 125–139.

Waters, C. K. (1994). Genes made molecular. *Philosophy of Science, 61*(2), 163–185.

Waters, C. K. (2004). What concept analysis in philosophy of science should be (and why competing philosophical analyses of gene concepts cannot be tested by polling scientists). *History and Philosophy of the Life Sciences, 26*(1), 29–58.

Waters, C. K. (2007). Causes that make a difference. *The Journal of Philosophy, 104*(11), 551–579.

Waters, C. K. (2008). Beyond theoretical reduction and layer-cake antireduction: How DNA retooled genetics and transformed biological practice. In M. Ruse (Ed.), *The Oxford handbook of the philosophy of biology* (pp. 238–262). Oxford: Oxford University Press.

Watson, J. D., & Crick, F. H. C. (1953). A structure for deoxyribose nucleic acid. *Nature, 171*, 737–738.

Weber, M. (1998). *Die Architektur der Synthese. Entstehung und Philosophie der modernen Evolutionstheorie*. Berlin: de Gruyter.

Weber, M. (2005). *Philosophy of experimental biology*. Cambridge: Cambridge University Press.

Weber, M. (2008). Critical notice: Darwinian reductionism. *Biology and Philosophy, 23*(1), 143–152.

Weinberg, R. A. (1998). *One renegade cell*. New York: Basic Books.

Wetzel, L. (2011). Types and tokens. In: E.N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2011 Edition). http://plato.stanford.edu/archives/spr2011/entries/types-tokens/

Wilson, D. S. (1988). Holism and reductionism in evolutionary ecology. *OIKOS, 53*(2), 269–273.

Wilson, J. M. (2010). Non-reductive physicalism and degrees of freedom. *British Journal for Philosophy of Science, 61*(2), 279–311.

Wimsatt, W. C. (1974). Complexity and organization. *PSA, 1972*, 67–86.

Wimsatt, W. C. (1976a). Reductive explanation: A functional account. *PSA, 1974*, 671–710.

Wimsatt, W. C. (1976b). Reductionism, levels of organization, and the mind-body problem. In G. G. Globus (Ed.), *Consciousness and the brain* (pp. 205–267). New York: Plenum Press.

Wimsatt, W. C. (1980). Reductionistic research strategies and their biases in the units of selection controversy. In T. Nickles (Ed.), *Scientific discovery: Case studies* (pp. 213–259). Dordrecht: Reidel.

Wimsatt, W. C. (1986). Forms of aggregativity. In: A. Donagan, A. N. Perovich Jr., & M. V. Wedin (Eds.), *Human nature and natural knowledge. Essays Presented to Marjorie Grene on the Occasion of Her Seventy-Fifth Birthday* (pp. 259–291). Boston: Reidel.

Wimsatt, W. C. (1997). Aggregativity: Reductive heuristics for finding emergence. *Philosophy of Science, 64*, 372–384.

Wimsatt, W. C. (2006a). Reductionism and its heuristics: Making methodological reductionism honest. *Synthese, 151*(3), 445–475.

Wimsatt, W. C. (2006b). Aggregate, composed, and evolved systems: Reductionistic heuristics as means to more holistic theories. *Biology and Philosophy, 21*(5), 667–702.

Wimsatt, W. C. (2007). *Re-engineering philosophy for limited beings. Piecewise approximations to reality*. Cambridge: Harvard University Press.

Wimsatt, W. C., & Sarkar, S. (2006). Reductionism. In S. Sarkar & J. Pfeifer (Eds.), *The philosophy of science: An encyclopedia* (pp. 696–702). New York: Routledge.

Winther, R. G. (2006). Parts and theories in compositional biology. *Biology and Philosophy, 21*, 471–199.

Winther, R. G. (2011). Part-whole science. *Synthese, 178*, 397–427.

Woodward, J. (2002). What is a mechanism? A counterfactual account. *Philosophy of Science, 69*(3), 366–377.

Woodward, J. (2003). *Making things happen. A theory of causal explanation*. Oxford: Oxford University Press.

Woodward, J. (2010). Causation in biology: Stability, specificity, and the choice of levels of explanation. *Biology and Philosophy, 25*(3), 287–318.

Woodward, J. (2011). Scientific explanation. In: E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2011 Edition). http://plato.stanford.edu/archives/win2011/entries/scientific-explanation/

Wright, C. D. (2012). Mechanistic explanation without the ontic conception. *European Journal of Philosophy of Science, 2*(3), 375–394.

Wright, C. D., & Bechtel, W. (2007). Mechanisms and psychological explanation. In P. Thagard (Ed.), *Philosophy of psychology and cognitive science* (pp. 31–79). Amsterdam: Elsevier.

Yoo, J. (2008). New hope for non-reductive pysicalism. In A. Hieke & H. Leitget (Eds.), *Papers of the 31st international Wittgenstein symposium: Reduction and elimination in philosophy and the sciences* (Vol. 16, pp. 408–411). Kirchberg am Wechsel: Eigner.

Index

M.I. Kaiser, *Reductive Explanation in the Biological Sciences*, History, Philosophy and Theory of the Life Sciences 16,
DOI 10.1007/978-3-319-25310-7