

Synthese Library 440

Studies in Epistemology, Logic, Methodology,
and Philosophy of Science

Quentin Ruyant

Modal Empiricism

Interpreting Science without Scientific
Realism



Springer

Synthese Library

Studies in Epistemology, Logic, Methodology,
and Philosophy of Science

Volume 440

Editor-in-Chief

Otávio Bueno, Department of Philosophy, University of Miami, USA

Editors

Berit Brogaard, University of Miami, USA

Anjan Chakravartty, University of Notre Dame, USA

Catarina Dutilh Novaes, VU Amsterdam, The Netherlands

Darrell P. Rowbottom, Lingnan University, Hong Kong

Emma Ruttkamp, University of South Africa, South Africa

Kristie Miller, University of Sydney, Australia

The aim of *Synthese Library* is to provide a forum for the best current work in the methodology and philosophy of science and in epistemology. A wide variety of different approaches have traditionally been represented in the Library, and every effort is made to maintain this variety, not for its own sake, but because we believe that there are many fruitful and illuminating approaches to the philosophy of science and related disciplines.

Special attention is paid to methodological studies which illustrate the interplay of empirical and philosophical viewpoints and to contributions to the formal (logical, set-theoretical, mathematical, information-theoretical, decision-theoretical, etc.) methodology of empirical sciences. Likewise, the applications of logical methods to epistemology as well as philosophically and methodologically relevant studies in logic are strongly encouraged. The emphasis on logic will be tempered by interest in the psychological, historical, and sociological aspects of science.

Besides monographs *Synthese Library* publishes thematically unified anthologies and edited volumes with a well-defined topical focus inside the aim and scope of the book series. The contributions in the volumes are expected to be focused and structurally organized in accordance with the central theme(s), and should be tied together by an extensive editorial introduction or set of introductions if the volume is divided into parts. An extensive bibliography and index are mandatory.

More information about this series at <http://www.springer.com/series/6607>

Quentin Ruyant

Modal Empiricism

Interpreting Science Without Scientific
Realism

 Springer

Quentin Ruyant
General Philosophy of Science
Universidad Nacional Autónoma de México
Mexico, Mexico

ISSN 0166-6991

ISSN 2542-8292 (electronic)

Synthese Library

ISBN 978-3-030-72348-4

ISBN 978-3-030-72349-1 (eBook)

<https://doi.org/10.1007/978-3-030-72349-1>

© Springer Nature Switzerland AG 2021

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This book aims at presenting the precise articulation of a pragmatist stance towards science, a stance that takes the form of an anti-realist position in the debate on scientific realism: modal empiricism.

Various forms of scientific realism seem to dominate the current philosophical landscape, whether they are actively defended or simply assumed for the sake of doing metaphysics. Such a “realist stance” towards science is often associated with bare common sense, and in contrast, anti-realism could seem like nothing but misplaced scepticism. Perhaps, some philosophers are even reluctant to call themselves anti-realist for the fear of being associated with people who entertain a certain defiance towards science, and admittedly, such defiance is a pressing problem of our times. However, entertaining overly optimistic positions about science might not be the best way of addressing these worries. But more importantly, as explained in Chap. 1 of this book, I think that it is a mistake to view anti-realism as a lack of trust in science, or to assume that realists actually side with scientists more than anti-realists do. What is really in question is the *interpretation* of science, its activities, aims and achievements, not trust in science. This is how modal empiricism should be understood, and I hope that this book will demonstrate the viability of this project.

Modal empiricism, as its name implies, is a version of empiricism that is committed to the idea that there are possibilities in the world and natural constraints on these possibilities. According to modal empiricism, our best scientific theories reflect the way these constraints affect our possible observations and actions and thus allow us to navigate in this world successfully. This is the kind of understanding that science provides. Abstract representations, such as scientific models, are generally indexical: they convey norms for their applications in particular contexts, and these norms are geared towards empirical success in all possible situations that are accessible to us. In this book, I argue that modal empirical adequacy is achievable, that modal empiricism makes better sense of scientific practice than constructive empiricism does, and that it can respond to the main arguments in the debate on scientific realism without assuming that theories are true descriptions of a mind-independent reality.

The developments and arguments of this book are largely based on an account of epistemic representation, presented at the end of Chap. 2 and detailed in Chap. 3, which acts as a framework for modal empiricism. This account of epistemic representation is meant to be an important contribution in its own right, and I hope that it will be considered as such. In his Stanford Encyclopedia entry dedicated to the structure of scientific theories, Winther distinguishes a syntactic, a semantic and a pragmatic view of theories, and he claims that “[t]he analytical framework of the Pragmatic View remains under construction”, implying that it is not as developed as its syntactic and semantic counterparts. This might be due to the emphasis on informal and practical aspects that characterises pragmatic approaches, as well as to the pluralist stance often adopted by pragmatist philosophers. Nevertheless, I believe that it is possible to make some steps in the direction of a more developed unified “analytical framework” for a pragmatic view of theories, without neglecting the complexities of scientific practice. My proposed reconciliation, which I call the “two-stage account of epistemic representation”, consists in understanding abstract, communal representations in terms of norms constraining or licencing the contextual use of representational vehicles. This view is largely inspired by Grice’s philosophy of language and, in particular, by the distinction between speaker meaning and expression meaning. Assuming this two-stage account, the debate on scientific realism can bear on the status of the representational norms developed by the scientific community and, in particular, on their aims and the implications of their success. In sum, as said earlier, this is a debate about interpreting science.

According to modal empiricism, the aim of science is empirical success for all possible contexts of use, whatever one’s purpose. The corresponding notion of empirical adequacy is developed in Chap. 4. I explain how it differs from van Fraassen’s notion, in particular, by its modal and situated character, and I argue that it is better apt to account for scientific rationality. A notion of situated possibility, together with an inductivist epistemology, is presented in Chap. 5, so as to defuse sceptical reluctance to endorse natural modalities. Chapters 6 and 7 address more traditional themes of the debate on scientific realism: the no-miracle argument and the pessimistic meta-induction. I explain how modal empiricism can account for scientific success without inference to the best explanation, how exactly it differs from structural realism and how these differences make it better apt to respond to the problem of theory change. Finally, Chap. 8 is concerned with semantic aspects. I explain how one can make sense of scientific discourse and even take it “at face value”, without being a scientific realist, by adopting a pragmatist conception of truth. This final chapter also presents a pragmatist and revisionist approach towards metaphysics that incorporates indexical and normative aspects at its core. Although this is not the most developed part of the book, it promises to deliver an important message: that philosophy, including metaphysics, can be (and then, should be) practically relevant.

The philosophical stance that emerges from this book consists in assuming that philosophy of science, even when it is interested in a broad picture or in metaphysical issues, should put situated representations of concrete objects at the centre stage of its analysis, rather than idealistic (and, in effect, non-existent)

representations of the whole universe. Modal empiricism is more than a mere position in the debate on scientific realism: it is a pragmatist framework for interpreting science and for doing philosophy of science.

I very much enjoyed writing this book, and I hope that its readers will enjoy reading it and gain a new perspective on science.

Mexico City, Mexico
October 2020

Quentin Ruyant

Acknowledgements

Many thanks to Otávio Bueno for giving me the opportunity to write this book; to the members of my PhD jury, Alexandre Guay, Pierre Joray, Filipe Drapeau-Contim, Peter Verdée and Anouk Barberousse, for their comments on my work; to the anonymous reviewer of this book, who helped me improve its content significantly; to the UNAM for financially supporting my research; to Stéphanie Brabant for her valuable help; and to Catherine Avery for her support and dedication.

Contents

- 1 The Debates on Scientific Realism** 1
 - 1.1 What Is at Stake With Scientific Realism?..... 1
 - 1.2 The Components of Scientific Realism 3
 - 1.3 Modal Empiricism: An Interpretation of Science 6
 - 1.4 The Two Battles to Be Fought..... 9
 - References 12
- 2 Theories, Models and Representation** 13
 - 2.1 Understanding Scientific Theories 13
 - 2.2 From Statements to Models 14
 - 2.2.1 The Statement View 14
 - 2.2.2 The Structuralist Semantic View 15
 - 2.2.3 The Moderate Semantic View 16
 - 2.3 From Semantics to Pragmatics 17
 - 2.3.1 The Model–Theory Relationship..... 18
 - 2.3.2 The Model–World Relationship..... 19
 - 2.3.3 The Model–Experience Relationship 21
 - 2.4 From Users to Communities..... 23
 - 2.4.1 User-Centred Accounts of Representation 24
 - 2.4.2 Anything Goes?..... 26
 - 2.4.3 Towards a Two-Stage Account of Representation..... 27
 - 2.5 A Tension Between Contextuality and Unity 29
 - References 30
- 3 Contextual Use and Communal Norms**..... 35
 - 3.1 The Two-Stage Account of Epistemic Representation 35
 - 3.2 First Stage: Contextual Use..... 37
 - 3.2.1 Contexts 37
 - 3.2.2 Interpretation..... 40
 - 3.2.3 Concrete and Abstract Interpreted Models 43
 - 3.2.4 Accuracy 45
 - 3.2.5 Model Choice 47

3.3	Second Stage: Communal Status	48
3.3.1	Two Senses of Representation	48
3.3.2	Licensing	49
3.3.3	General Models and Indexicality	50
3.3.4	Relevance	53
3.3.5	Modularity and Composition	56
3.3.6	Explanations	57
3.4	The Norms of Representation in Science	59
3.4.1	Theoretical Unity	60
3.4.2	The Interpretation of Theoretical Terms	63
3.4.3	Layers of Representation	66
3.5	Epistemic Values and the Axiological Debate	68
	References	71
4	Modal Empirical Adequacy	73
4.1	Empirical Adequacy as an Axiological Notion	73
4.2	From Contextual Accuracy to General Adequacy	76
4.2.1	An Ampliative Notion	76
4.2.2	Situations and Possible Contexts	77
4.2.3	Different Versions of Empirical Adequacy	80
4.2.4	Are Modalities Innocuous?	82
4.2.5	Exotic Possible Situations	85
4.2.6	Modal Empirical Adequacy	88
4.3	From Models to Theories	90
4.3.1	Two Options	90
4.3.2	Unproblematic Failure as Inaccuracy	92
4.3.3	Unproblematic Failure as Irrelevance	94
4.3.4	Theory Evolution and Theory Change	98
4.4	Comparison with van Fraassen's Account	100
4.4.1	Situations Versus the Universe	101
4.4.2	Relevance Norms Versus Observables	103
4.4.3	Modality Versus Extensionality	105
4.5	Modal Empiricism as Pragmatism	108
	References	109
5	Situated Possibilities, Induction and Necessity	111
5.1	The Sceptical Challenge	111
5.2	Situated Possibilities	113
5.2.1	Conceivable and Possible Situations	114
5.2.2	Statements of Necessity	117
5.2.3	Why Should We Accept Situated Possibilities?	121
5.3	The Inductive Route Towards Necessity	124
5.3.1	Induction and Underdetermination	125
5.3.2	The Inductive Route	127

5.3.3	Modal Underdetermination	131
5.3.4	Modal Statements are Not Underdetermined	133
5.3.5	Modal Conflicts in Scientific Practice	136
5.4	Laws of Nature	139
5.4.1	Induction and Projectibility	139
5.4.2	Laws and Accidents	142
5.5	Do We Need More?	144
	References	145
6	Scientific Success	149
6.1	The Miraculous Success of Science	149
6.2	The No-Miracle Argument	152
6.2.1	Underdetermination and Non-Empirical Virtues	152
6.2.2	Against Inference to the Best Explanation	156
6.2.3	The No-Miracle Argument	160
6.3	An Inductivist Response	164
6.3.1	The Selectionist Explanation	164
6.3.2	Induction on Contexts	166
6.3.3	Induction on Models	169
6.4	Objections to the Inductivist Response	171
6.4.1	Sample, Population and Modalities	171
6.4.2	The Base-Rate Fallacy and Fallibilism	174
6.4.3	Uniformity of Nature and Relativity	176
6.5	How Far Are We From Realism?	179
	References	180
7	Theory Change	183
7.1	The Pessimistic Meta-Induction	183
7.2	Structural Realism and Newman's Objection	187
7.2.1	The Objections Against Structural Realism	187
7.2.2	How to Escape Newman's Objection?	189
7.2.3	Transposition to the Semantic View	191
7.3	"Real" Relations and Theory Change	193
7.3.1	Newman's Objection in Modal Logic	193
7.3.2	Are Modal Relations Real?	194
7.3.3	Which Modal Relations Are Retained in Theory Change?	197
7.4	Relativity and Fundamentality	199
7.4.1	Real Patterns and Locators	200
7.4.2	Fallibilism	202
7.5	Modal Empiricism: The Best of Both Worlds	203
	References	205
8	Semantic Pragmatism	209
8.1	Anti-Realism, Acceptance and Belief	209
8.2	Correspondence and Pragmatic Truth	212
8.3	Why Semantic Realism?	214

8.4 A Pragmatist Alternative.....	218
8.5 Scientific Objectivity.....	221
8.6 Modal Empiricism and the Role of Metaphysics	223
References	229

Chapter 1

The Debates on Scientific Realism



Abstract The debate on scientific realism results from a tension between the empiricist methodology, which is a defining feature of science, and claims to the effect that science can unveil the fundamental nature of reality. What distinguishes realist and anti-realist positions is not necessarily that the former take scientific knowledge “at face value” or take the side of scientists in general while the latter do not. Rather, realists and anti-realists propose different ways of interpreting science as a whole, and in particular its aim (axiological realism), its possible achievements (epistemic realism) and its content (semantic realism). The aim of this book is to defend an interpretation that potentially applies to each of these three levels: modal empiricism. This position purports to be the articulation of a pragmatist stance towards science. This introductory chapter briefly presents the position, then outlines the structure of the book.

1.1 What Is at Stake With Scientific Realism?

What do we know about reality? When asking this question not about reality in general, but about a specific subject area, the natural attitude is to turn towards science for an answer. What do we know about living organisms? At least what biology tells us: complex organisms are composed of living cells, which reproduce by duplicating their genetic code, which is stored in DNA molecules, and so on. What do we know about combustion phenomena? Well, what we can read in chemistry books: they involve reactions where big molecules break into smaller ones and release energy.

Now turning back to our initial, general question, its answer could be that our best knowledge about reality is provided by the best theories offered by science. The content of these theories should be taken *at face value*: they literally describe what exists in the world, and the entities described (objects, processes, properties, relations...) do not depend on our interests, representations or activities for their existence. Science unveils the nature of reality. This is, roughly, the content of the doctrine of scientific realism.

It could seem, on the surface of it, that scientific realism is nothing but a commonsense trust in the capacity of science to give us knowledge about the world. To those who would object that blind trust is irrational, the realist can respond by qualifying her attitude: only mature science is concerned, for example. But a general distrust that would concern all of science could be perceived as misplaced. This could explain why realism sometimes seems to be used in the philosophical literature as a mark of seriousness, and why it is often deemed important that such or such philosophical view is “compatible with scientific realism”.

I think that such a surface reading of the debate, which places anti-realism in an uncomfortable position, is inaccurate. This is not to deny that some anti-realist positions entertain an attitude of defiance towards science in general, but this is not an essential feature of anti-realism. Anti-realists rarely deny that science is extraordinarily successful, that it has progressed, and that theories can explain a variety of phenomena. An anti-realist can even accept that science gives us knowledge that is more valuable than the one that we could obtain by other means, and she can make sense of the natural attitude consisting in turning to scientific theories for answers to specific questions about the world. An anti-realist does not necessarily put into question the achievements of science or the content of theories, but rather challenges their realist interpretation. Is science really concerned with the deep nature of reality? Or is it concerned with something more practical and mundane? Science provides explanations and understanding, but is there really more to explanations and understanding than the functions they play in our interactions with the world? So, the choice is not really (or not always) between trust and scepticism towards scientific knowledge, but rather between various ways of interpreting this knowledge. In sum, I believe that the debate on scientific realism is more accurately framed as a debate over the interpretation of science.

The aim of this book is to present and defend a certain interpretation of science as an alternative to scientific realism. I call the corresponding position “modal empiricism”. Empiricist positions are often characterised as the idea that the aim of scientific theories is to “save the phenomena”, and as a first approach, modal empiricism can be understood as the idea that theories aim at saving the *possible* phenomena. I will defend this position by arguing that it fares better than its opponents in the debate on scientific realism. However, in light of what has just been said, this position should not be merely understood as a technical solution in response to specific philosophical arguments, nor as a mere form of scepticism towards science. My ambition is to propose a positive way of interpreting science, which has implications beyond the debate on scientific realism, including, for example, for the metaphysics of science. More precisely, modal empiricism purports to be the articulation of a pragmatist stance towards science, which puts practice at the centre of interpretational issues, and I am convinced that this position constitutes the best way of articulating this pragmatist stance.

Interpreting science can mean interpreting the activities of scientists and the aims of these activities, or it can mean interpreting the products of these activities, in particular, scientific theories. For this reason, the debate is multifaceted, which can easily be overlooked when one looks at it in terms of defiance versus trust. Before saying more about modal empiricism, let us examine these multiple facets.

1.2 The Components of Scientific Realism

Scientific realism is often described as a commitment to three theses: the metaphysical thesis, according to which a mind-independent reality exists, a semantic thesis, according to which scientific theories are “about” this reality, and an epistemic thesis, according to which scientific theories are at least approximately true. Let us detail them in more precise language.

The metaphysical thesis has two components. First, there is the idea that an external, mind-independent reality exists, which opposes idealism, according to which reality is mental. Second, there is the idea that reality is structured in a way that is in principle intelligible, which typically opposes both Kantian views and some versions of constructivism, according to which the phenomena that our intellect can grasp are somehow constituted by our representations or activities. According to the metaphysical realist, “the world has a definite and mind-independent natural-kind structure” (Psillos 1999, p. xvii).

None of these components of realism will be discussed at length in this book, because a metaphysical stance is often made implicit by the semantic stance one adopts, or at least, it seems reasonable to be clear on semantic issues before discussing metaphysical issues: after all, we need language to talk about reality. Nevertheless, as I said earlier, modal empiricism has implication for metaphysics, and this topic will be touched upon in the concluding chapter of this book.

The semantic thesis of scientific realism has to do with the relationship between our representations and reality. It consists in adopting *semantic realism*. Again, two components can be distinguished in the realist stance: (1) a truth-conditional semantics, and (2) a conception of truth that is not epistemically constrained (Shalkowski 1995) (or that is such that truth conditions are “potentially evidence transcendent” Miller 2003). Such a conception of truth can be, for example, the idea that a statement is true if it corresponds to reality.

The first component, truth-conditional semantics, differentiates semantic realism from what Psillos (1999) calls “eliminative instrumentalism”, which takes theories to be mere instruments without truth-values (theories would be good or bad rather than true or false). The second component, the conception of truth, differentiate it from “reductive empiricism”. Although it accepts that theories have truth-values, reductive empiricism attempts to reinterpret the content of scientific theories in terms of mere observables (theories would not be about mind-independent entities), which makes theoretical truth epistemically constrained. In sum, semantic realism claims that theories are capable of being true or false, in virtue of reality. Or, as Psillos puts it:

The semantic stance takes scientific theories at face-value [...]. Theoretical assertions are not reducible to claims about the behaviour of observables, nor are they merely instrumental devices for establishing connections between observables. The theoretical terms featuring in theories have putative factual reference. So, if scientific theories are true, the unobservable entities they posit populate the world.

Here, Psillos puts emphasis on observables, so as to contrast semantic realism with reductive empiricism, but semantic realism is more general. A conventionalist who claims that scientific theories are nothing but implicit definitions for a theoretical vocabulary, and that therefore they are true by convention, is not a semantic realist, for instance. For a semantic realist, the content of scientific theories should not be interpreted as mere conventions, nor in terms of notions such as measurements, observations, intentions, information, social norms or any other epistemically loaded or anthropocentric term, at least not if these are taken to be unanalysable, irreducible notions associated with the users of the theory. Such notions are incompatible with the idea that scientific theories describe a mind-independent reality. This idea seems to be an important desideratum in the metaphysics of science (Bell 2004's disdain for the notion of measurement is often cited to that effect in the philosophy of physics literature). Not all positions called "realism" satisfy this condition. For example, Putnam's internal realism explicitly rejects semantic realism. Nevertheless, semantic realism is generally accepted as an essential component of scientific realism.

The fact that scientific realism incorporates a semantic component should clarify why the debate on scientific realism can be understood as a debate over how science ought to be interpreted, and why an anti-realist does not necessarily deny that science gives us knowledge. Maybe it is hard to see how a theory conceived of as a mere instrument for making predictions could give us knowledge. Perhaps it is more accurate to say that an instrument *affords* practical knowledge, or "know-how", because one can learn how to use it as a tool for various purposes, but it seems to be part of common-sense intuitions that science provides *factual* knowledge, or "know-that". An eliminative instrumentalist would most certainly reject these intuitions. However, this is not something that a reductive empiricist would deny. The reductive empiricist would only reinterpret this knowledge in terms of observables.

The term "reinterpret" as well as its opposite "at face value", employed by Psillos, suggest that the semantic realist pays more respect to science than the semantic anti-realist does. I think that this is a mistake. Close attention to the way theories are used to represent shows that this idea of interpreting theories "at face value" is far from clear. As will be explained in Chap. 2, scientific theories are rarely considered to be linguistic statements by contemporary philosophers, so there is work to do in order to understand what theoretical truth amounts to exactly, and the realist route, which would take truth not to be epistemically constrained, is not necessarily the most natural one. At any rate, no semantic theory is imposed on us by scientific discourse, so there is no reason to think that the realist stands on the side of scientists in these matters.

As will be explained shortly, modal empiricism can be understood as a position that challenges semantic realism, but not in the same way as a reductive empiricist does. The debate on semantic realism will be addressed explicitly in the conclusion of this book. It will also appear more or less implicitly in many discussions along the way. However, this work is primarily focused on the debate over the epistemic thesis of scientific realism. One reason for this focus is that this epistemic thesis takes centre stage in contemporary discussions, while the semantic thesis is rarely

discussed. Discussions on the epistemic thesis generally take semantic realism for granted (and I will do so as well, for the sake of the discussion, until the concluding chapter). Another reason is that just as it is more appropriate to discuss metaphysical issues after having clarified semantic issues, I think that it makes more sense to discuss semantic issues after having clarified epistemic ones.

So, let us now turn to the last component of scientific realism: the epistemic thesis. According to this thesis, science is able to produce true, or approximately true theories (in a realist sense of true). To quote Psillos again, this entails that “the entities posited by them, or, at any rate, entities very similar to those posited, do inhabit the world”. This thesis is more easily associated with an attitude of trust, as opposed to defiance, towards science, since it concerns its achievements. However, things are a bit more complex than this.

It can be useful to break down this epistemic thesis into two parts: first, the idea that science *aims at truth*, and secondly, the idea that it is successful in this aim. Let us call the first aspect *axiological realism* and the second one *epistemic realism*.

The advantage of this formulation is that it makes room for van Fraassen’s approach towards the debate on scientific realism. Van Fraassen (1980, p. 8) characterises realism as the claim that “[s]cience aims to give us, in its theories, a literally true story of what the world is like; and acceptance of a scientific theory involves the belief that it is true”. He argues for a different axiological position, constructive empiricism, according to which the aim of science is to produce empirically adequate theories.

As proposed by Ghins (2017), the difference between axiological realism and epistemic realism can also be understood in terms of a distinction between a “pragmatic approach”, which is focused on scientific activity, and a “contemplative approach”, focused on the end-products of this activity: scientific theories. These two approaches imply different kinds of discussions. A pragmatic approach demands that we pay close attention to scientific practice. The purpose of an axiological thesis is to make sense of scientific practice as a rational activity, or to inquire into the aims and norms of this activity. For the realist, the aim of science is truth, while for the empiricist, it is empirical adequacy: conformity with experience is the final judge for the acceptability of theories and hypotheses. Others, which we shall call pluralists, deny that science has a unified aim. These are stances that concern the pragmatic approach. On the contrary, a contemplative approach will be focused not on the activity itself, but on the theories that this activity produces. An epistemic thesis is concerned with the justification of various attitudes towards these theories. Issues of justification can be addressed more abstractly, without paying too much attention to the way theories are actually constructed. However, the two are not unrelated, since it would be odd to attribute an aim to science if this aim was not achievable, or to consider that science achieves something that is not even part of its aims (although Lyons 2005 has defended the former).

Here, we can see a second reason why anti-realism should not necessarily be considered an attitude of defiance towards sciences: if the aim of science is not truth, then one can be an anti-realist and at the same time consider that science is a successful endeavour.

The position that I will defend in this book, modal empiricism, challenges both axiological and epistemic realism. According to modal empiricism, the aim of science is not truth, but modal empirical adequacy, and science is successful in this aim, in the sense that the adequacy of theories can be justified by experience. The axiological thesis is roughly the focus of the first half of this book (from Chaps. 2 to 4), and the epistemic thesis is defended in the second half (from Chaps. 5 to 7). Let us now present the main characteristics of this position in more detail.

1.3 Modal Empiricism: An Interpretation of Science

Why challenge scientific realism? There are many reasons, but perhaps the main one is that there is an inherent tension in the position. This tension is introduced by the semantic component, in particular the assumption that truth is not epistemically constrained, and it is manifest in the epistemic component, and in particular in the idea that science can achieve truth. In a word, semantic realism introduces a principled gap between truth and our epistemic capacities, and this casts doubt on the idea that truth, as understood by the realist, is achievable.

The tension is particularly salient in the case of science, because arguably, systematic confrontation with experience (rather than blind reliance on intuitions, for instance) is the cornerstone of scientific methodology. But how can we claim to have knowledge of universal laws of nature if those are only ever confirmed by particular, limited sets of observations? And how can we claim to have knowledge of unobservable entities such as electrons if their existence is only ever confirmed by their indirect consequences on our observations? What is at stake is the validity of ampliative modes of inference, that is, inferences that go beyond mere appearances.

All this could seem like unfounded philosophical worries to the scientific mind: “Look how well our theories work, are you doubting that electrons or gravity really exist?” To which the anti-realist could respond: “Look at the history of science. Theories come and go. Being willing to abandon theories in the face of new evidence is at the core of scientific methodology. Are you really confident that electrons and gravity will never be replaced by different concepts?” This is, in a nutshell, how the contemporary debate is framed.

As said earlier, the aim of this book is to present and defend an original position in this debate, which I call “modal empiricism”. The term “empiricism” marks the emphasis of the position on our interaction with the world through experience, rather than on reality itself, and the term “modal” refers to the notions of possibility and necessity. As already explained, modal empiricism can be roughly understood as

the idea that scientific theories “save the possible phenomena”.¹ The position can be summarised as follows:

Modal empiricism (1): the aim of science is to produce theories that correctly account, in a unified way, for all our possible manipulations and observations within particular domains of experience, and science is generally successful in this aim.

This position challenges the epistemic and axiological thesis of scientific realism presented above. However, it is still committed to the idea that science has a certain unity, and that it can be characterised by an aim, which is empirical adequacy. In this sense, it shares similarities with van Fraassen’s constructive empiricism, which also challenges the axiological thesis, and considers that empirical adequacy is the aim of science. The main difference lies in the modal aspect of modal empiricism, which makes the position more ambitious than traditional versions of empiricism. This aspect should be understood in terms of *natural* modalities, that is, not in terms of what is or is not conceivable or likely to be the case, but in terms of what is or is not possible in this world given natural constraints on phenomena.

In the conclusion of this work, I will suggest a possible reformulation of the position as challenging the semantic thesis of scientific realism instead of the epistemic and axiological ones. Adopting a different, pragmatist notion of theoretical truth, which happens to coincide with the modal version of empirical adequacy just presented, it is possible to maintain that indeed, science aims at truth and that it is successful in this respect. However, this is not the notion of truth that the scientific realist has in mind. This reformulation is a way of making sense of the natural attitude described in the opening of this chapter, without troubling ourselves with the unnecessary features of scientific realism, and as I will argue in the conclusion of this book, this semantics is in line with Peirce’s conception of pragmatism, and it is also consistent with scientific discourse. This gives us another understanding of modal empiricism:

¹The label “modal empiricism” has been co-opted in the epistemology of metaphysics to denote a different position from the one that I have in mind. It is employed in particular in debates concerning the validity of an inference from conceivability to metaphysical possibility (for example, from the conceivability of “phenomenal zombies” to their metaphysical possibility in the philosophy of mind). In this context, modal empiricism opposes modal rationalism. The idea can be that metaphysical possibility should be understood in terms of conceivable empirical justification (Hanrahan 2009), or that knowledge about metaphysical possibilities can be acquired by experience (Roca-Royes 2017). Although there are a few similarities between these positions and the one defended in this book, I use “modal empiricism” in a different sense: the position to which I refer is not concerned with metaphysical modalities, but with natural modalities, and it does not oppose modal rationalism in the epistemology of metaphysics, but non-modal empiricism in the philosophy of science. Giere (Churchland and Hooker 1985, ch. 4) and Ladyman and Ross (2007) use the term in the same sense as I do. I hope that the fact that its different meaning is used in another area of philosophy will be enough to avoid confusion.

Modal empiricism (2): our best scientific theories are true, or approximately true, in the sense that they correctly account, in a unified way, for all our possible manipulations and observations within their domains.

The idea is to identify truth with a certain notion of ideal success. Assuming this notion of truth, modal empiricism can be understood as a form of realism. However, this is only true if realism is understood in a broad sense, and not in the traditional sense of the term. Modal empiricism is closer to Putnam's internal realism than it is to standard scientific realism. I will argue that this approach has the capacity to connect metaphysical debates to more tangible considerations.

As its formulation in terms of pragmatic truth makes clear, modal empiricism is a form of pragmatism, in the tradition of Peirce (with whom I also share a commitment to natural modalities). I understand pragmatism as being characterised by an emphasis on a practical, active conception of knowledge, as opposed to what Dewey calls the "spectator conception of knowledge". From a pragmatist perspective, the starting and end point of any inquiry into the aim and achievements of science should be scientific practice. In order to interpret scientific theories, we should first have a look at how they are used in particular contexts, and an interpretation should have practical implications in one way or another.

Modal empiricism is based on an account of scientific representation that clearly articulates the contextual uses of theories and the abstract structure of these theories. In a pragmatist spirit, I understand scientific theories and models as conveying norms of representation that constrain particular uses. This account also does justice to the fact that in general these contextual uses are not passive, but partly performative, hence the mention of manipulations in the definition above. This account of scientific representation constitutes our framework for inquiring into the aims and achievements of science.

According to modal empiricism, the aim of science is to produce norms of representation that are ideally empirically successful, in all circumstances, whatever one's particular purpose is. As I just said, these norms are conveyed by scientific theories and models, so the aim of science is to produce ideally successful models and theories. This notion of ideal success is not characterised in abstract terms, for example, in terms of a general correspondence between the structure of the theory and the world, but rather by quantifying over possible uses of the theory. This is the situated aspect of the position. Combined with the idea that experimentation is not a passive activity, this leads us quite directly to the modal aspect of the position. Merely possible observations could concern actual states of affairs, but merely possible manipulations cannot, because they would create non-actual states of affairs. So, we have to assume that ideal success concerns mere possibilities. An adequate theory tells us what it is possible to do and observe in this world.

However, the kind of modality that characterises modal empiricism is quite distinct from traditional construals of natural modalities, for example in terms of laws of nature, precisely because it is situated. Understanding these modalities in terms of possible worlds is inappropriate, because these possibilities are, so to speak, anchored to actual contexts. I will propose an understanding in terms of possible situations.

This understanding of natural modality is crucial. It is what makes the position distinctively empiricist, as opposed to structural realism, for example. Modal empiricism does not claim that we can have knowledge of the laws of nature, or of the “modal structure of the world”, however it is interpreted. We can only have knowledge of the way natural constraints affect our observations and manipulations in context. Since the contexts to which we have access are limited by our situation in the universe and by our cognitive constitution, a form of epistemic relativity (or perspectivity) ensues. So, even though modal empiricism is more ambitious than non-modal versions of empiricism, it cannot be classified as a version of scientific realism. As I will argue in this book, this understanding of modality is what allows modal empiricism to be the best compromise in the debate on scientific realism, because it is able to make sense of scientific success without relying on problematic modes of inference, and it is not threatened by arguments based on theory change.

In sum, modal empiricism purports to articulate a pragmatist stance towards science, and this stance is characterised by the following aspects:

Normativity:	abstract representations convey norms that apply to concrete uses
Situatedness:	concrete uses are sensitive to a local context
Performativity:	concrete uses involve observations as well as manipulations
Modality:	ideal success also concerns merely possible uses

1.4 The Two Battles to Be Fought

Now that we have a better idea of what modal empiricism is, let us outline the structure of this book. The book can be roughly divided into three parts of two chapters each.

The first two chapters set the stage for the discussion by presenting the account of scientific representation on which modal empiricism is based. Chapter 2 is a review of the debates on the nature of scientific theories and models and their relations to users, experience and the world. This review serves as a motivation for an approach that takes as a starting point contextual, situated uses of scientific theories, while also paying attention to the communal, normative aspects of representation. An account of scientific representation and of the nature of scientific theories that takes this starting point is proposed in Chap. 3. As explained earlier, this account of scientific representation is the framework on which modal empiricism is based. According to this account, the main function of scientific models is to convey indexical norms of representation, and using a model to represent a concrete object in a particular context is applying these norms. This chapter introduces important notions, such as context, model, interpretation, relevance and accuracy, which are used throughout the rest of the book.

The remaining chapters are dedicated to the presentation and defence of modal empiricism. As a middle position between scientific realism and traditional versions of empiricism, modal empiricism must be defended against two main opponents:

the more pessimistic ones and the more optimistic ones. There are two battles to be fought.

The second part of the book, constituted of Chaps. 4 and 5, serves two roles: developing the position, and explaining why it should be adopted rather than other versions of empiricism. This is the first battle.

In Chap. 4, the focus is on the axiological component of the debate. According to empiricism, the aim of science is to produce theories that are empirically adequate, but various versions of empiricism can differ in their understanding of empirical adequacy. The chapter first develops the notion of empirical adequacy that modal empiricism endorses, with its characteristic modal and situated aspects, by adopting a bottom-up approach, starting from a consideration of success in contextual uses of models and leading up to a definition of ideal success at the theory level. This gives us the main criteria by which, according to modal empiricism, models and theories are accepted in science. Theoretical unification plays an important role in this respect. Then, the resulting position is compared to van Fraassen's constructive empiricism, the leading contemporary empiricist position. I argue that modal empiricism, thanks to its modal and situated aspects, better accounts for scientific practice as a rational activity, and in particular for the interventionist component of experimentation.

In Chap. 5, the focus is on modalities more specifically, and I start to focus on justification issues. The defining characteristics of situated modalities, conceived of in terms of alternative ways actual situations could be, are presented and discussed, as well as their differences with other kinds of modalities, in particular, the ones associated with the laws of nature. I give a few preliminary reasons to accept them in our ontology. Next, I address what I take to be the main obstacle to accepting modalities in an empiricist's world-view: the alleged principled impossibility of modal knowledge. In order to overcome this obstacle without resorting to problematic modes of inference, I propose an inductivist epistemology for situated modalities, based on the idea that realised possibilities are representative of non-realised ones. I show that since we do not know a priori which possibilities are realised and which are not, the status of statements of necessity is actually the same as the status of universal generalisations within a possible situation framework. This eliminates the reluctance one could have to adopt modal empiricism instead of non-modal versions of empiricism.

So much for the first battle. The second battle, which is the focus of the final part of the book, is against scientific realism and other related positions, such as structural realism. I take on the two arguments that structure the contemporary debate on scientific realism: the no-miracle argument for scientific realism in Chap. 6, and the pessimistic meta-induction in Chap. 7.

The no-miracle argument is an inference from empirical success to theoretical truth, and it involves a particular mode of inference: an inference to the best explanation. In Chap. 6, I examine this type of inference, the way it is used for solving underdetermination problems, and the criticisms that it has received. According to the realist, the main problem with empiricism is that it is unable to explain the success of science, and in particular the capacity of theories to lead to successful

novel predictions. In response, I explain how modal empiricism can actually account for novel predictions, in so far as the notion of empirical adequacy that it adopts can be justified. Novel predictions are among the possibilities accounted for by modal empirical adequacy, so, if our theories are empirically adequate, then these novel predictions are no miracle. I argue that the empirical adequacy of theories can indeed be justified by using a particular form of induction: an induction on the models of a theory.

The pessimistic meta-induction is a direct response to the no-miracle argument, which starts from the observation that most successful theories of the past have by now been replaced by better ones. Since these past theories are now considered false, this casts doubt on the idea that contemporary theories are true. A move in response to this argument consists in adopting structural realism. This position restricts realist claims to the *structure* of reality, without assuming that scientific theories correctly describe the *nature* of reality, or assuming that this nature is entirely structural. The argument that structural realists put forth is that there is structural continuity between successive theories, so that theory changes do not threaten this form of realism.

Some contemporary structural realists claim that the structure of reality is modal. This position is not far from modal empiricism, to the extent that some have claimed that a modal version of empiricism is nothing but structural realism. In Chap. 7, I examine this claim, and argue that it is untrue. The crucial difference between the two positions lies in the kinds of modalities that they adopt. The modal relations postulated by past theories are generally viewed as relative rather than absolute in light of their successors. I argue that for this reason, structural realism is actually unable to respond to the pessimistic meta-induction, unless it adopts a notion of relative modality, such as the situated modalities of modal empiricism. However, in such a case, the position cannot be considered a version of realism. This chapter clarifies the main differences between modal empiricism and scientific realism.

This argument completes our defence of modal empiricism against its opponents. Modal empiricism can make better sense of scientific practice as a rational activity than other versions of empiricism. The reasons to be a modal sceptic are unwarranted, given that statements of situated necessity are justified by induction as much as universal regularities are. Finally, modal empiricism can account for the empirical success of science, and it is not threatened by an induction on past theory changes. It comes out as the best position of compromise in the epistemic debate on scientific realism.

The conclusion of this book, Chap. 8, is an opportunity to return to the facets of the debate that have been left out in other chapters: the semantic and the metaphysical aspects. Semantic realism is rarely discussed in the literature, but why should we take it for granted? Does it make better sense of scientific discourse in general? Are there reasons from philosophy of language to adopt it with respect to scientific theories? I answer by the negative to these two questions: semantic realism is much more problematic than generally thought, and we have good reasons to assume that modal empirical adequacy is the right notion for delineating the cognitive content of scientific theories. This means adopting a pragmatist notion

of truth for theories, which comes very close to Peirce's conception of truth. In this sense (and this is a final twist), modal empiricism can actually be understood as a realist position, albeit not in the traditional sense.

This final move has important implications for the metaphysics of science. It favours adopting a revisionary pragmatist stance towards metaphysics, by connecting metaphysical considerations to more tangible aspects, and this has the potential of resolving (or perhaps sometimes dissolving) various debates through a pragmatist reinterpretation of them. At the same time, the commitment to natural modalities that characterises modal empiricism is a way not to fall back to an impoverished metaphysics that would fail to do justice to the various debates. For this reason, the position I advocate in this book should not be merely understood as yet-another-position-in-the-debate-on-scientific-realism, but it should be understood as a more far-reaching proposal. It is a proposal concerning the way of interpreting science in general, and the way of understanding the role of metaphysics of science. This proposal has a normative dimension: sound metaphysical speculations should seek systematic connections with possible experiences in order to have practical relevance. In sum, modal empiricism purports to be the precise articulation of a pragmatist approach towards the philosophy of science. I hope that it can serve as a useful basis for all philosophers sharing this pragmatist stance.

References

- Bell, J. (2004). Against measurement. In *Speakable and unspeakable in quantum mechanics* (pp. 213–231). Cambridge: Cambridge University Press.
- Churchland, P., & Hooker, C. (1985). *Images of science: Essays on realism and empiricism*. Chicago: University of Chicago Press.
- Ghins, M. (2017). Defending scientific realism without relying on inference to the best explanation. *Axiomathes*, 27(6), 635–651.
- Hanrahan, R. (2009). Consciousness and modal empiricism. *Philosophia*, 37(2), 281–306.
- Ladyman, J., & Ross, D. (2007). *Every thing must go: Metaphysics naturalized*. Oxford: Oxford University Press.
- Lyons, T. D. (2005). Toward a purely axiological scientific realism. *Erkenntnis*, 63(2), 167–204.
- Miller, A. (2003). The significance of semantic realism. *Synthese*, 136(2), 191–217. <https://doi.org/10.1023/A:1024742007683>
- Psillos, S. (1999). Scientific realism: How science tracks truth. In *Philosophical issues in science*. London: Routledge.
- Roca-Royes, S. (2017). Similarity and possibility: an epistemology of de re possibility for concrete entities. In B. Fischer & F. Leon (Eds.), *Modal epistemology after rationalism* (pp. 221–245). Cham: Springer. https://doi.org/10.1007/978-3-319-44309-6_12
- Shalkowski, S. A. (1995). Semantic realism. *Review of Metaphysics*, 48(3), 511–538.
- van Fraassen, B. (1980). *The scientific image*. Oxford: Oxford University Press.

Chapter 2

Theories, Models and Representation



Abstract Before discussing what the aim of science could be and whether it is achievable, it is important to be clear on what scientific theories are and how they are used to represent the world. This chapter proposes a synthesis of the literature on the subject, in particular the motivations for model-based approaches and the pragmatic and communal components of representation. The conclusion of this review is that a good conception of scientific representation must be model-based, and that it must reconcile the contextuality of experimentation and the unificatory power of scientific theories.

2.1 Understanding Scientific Theories

Before inquiring into the aim and achievements of science, it is worth being clear on the subject matter and, in particular, on the right way of understanding the nature of one the main products of scientific activity, the one upon which the debate on scientific realism primarily bears: scientific theories.

The present work rests on a model-based conception of scientific theories that incorporates strong pragmatic elements. A proper account of the relations between models and experience on the one hand, and models and theories on the other, is crucial for understanding the position that I will defend. This account will be laid out in the next chapter.

However, let us first define the stakes by briefly recalling the main motivations for model-based approaches, how these approaches have been developed by various authors following structuralist or pragmatist lines, and the debate on scientific representation that they have initiated. Three main desiderata for a good account of scientific theories will emerge from this presentation: a focus on models, a focus on users and a focus on communal norms.

2.2 From Statements to Models

It was once commonplace to think of scientific theories as linguistic entities. In the tradition of logical empiricism, and in particular in Carnap (1966)'s work, a theory was typically thought of as a deductively closed set of statements, including axioms expressed in a theoretical vocabulary, correspondence rules linking the theoretical and observable vocabulary, and their logical consequences. Let us call this the *statement view of theories*. This view has been challenged during the second half of the twentieth century, and today, theories are more often conceived of as families of models. Let us present the reasons for this move.

2.2.1 The Statement View

According to the statement view of theories described above, scientific representation is just one mode of linguistic representation. Scientific models do not play a prominent role. Several passages in the writings of Carnap suggest that the interest of scientific models is merely heuristic, psychological or pedagogical—a view which was not uncommon in scientific circles at the end of nineteenth century (Bailer-Jones 2013, ch. 2.2, 4.2). According to Carnap, models belong to the context of discovery, and they become superfluous once we are in possession of an axiomatic formulation of a theory.

Notwithstanding these depreciative views on scientific models, the statement view of theories does not preclude using logical models for meta-linguistic analysis of the content of theories. In Tarski's model theory, a logical model is a set-theoretical structure that is mapped to a vocabulary. It consists in a domain of objects and a valuation, such that the extensions of proper names and predicates of a language are specified in this domain (single objects for proper names, sets of n-tuples of objects for n-ary predicates). A logical model can satisfy linguistic statements according to the rules provided by Tarski. It acts as a truth-maker for the theory, and we can say that a theory is true if the world can be represented by a model that satisfies it. In the statement view, model theory is thus a way of analysing the world-theory relationship, but a scientific theory is still constituted of statements.

As is well-known, the statement view of theories has been challenged for various reasons during the second half of the twentieth century. Its main difficulties concerned notably the status of the correspondence rules connecting theories and observations, which encompass heterogeneous aspects, such as experimental techniques, linguistic meaning or domain-specific assumptions, and the dubious semantic distinction between theoretical and observational vocabulary on which the view rests. Other issues have to do with the fact that theories can receive more than one formulation and that they need not be axiomatised. Note that the statement view was partly normative: it was aiming at reconstructing scientific knowledge in a rigorous manner by extracting its logical features. Its aim was not to give

a descriptive account of science. Nevertheless, it became seen as overly abstract and disconnected from scientific practice. In particular, it became increasingly recognised that scientific models, and not general linguistic statements, are the main representational units in science (see Lutz 2017, for a review of these debates).

2.2.2 *The Structuralist Semantic View*

In an attempt to bring philosophical analysis closer to actual scientific practice, it has been proposed that scientific theories should be characterised as collections or families of models rather than as linguistic statements. This so-called *semantic view of theories*, originally proposed by Suppes (1960), has now become commonplace (Ladyman and Ross 2007; Giere 1997; Suppe 1989; van Fraassen 1980). A similar structuralist programme has been developed by Sneed (1971) and Stegmüller (1979).

In the semantic view, theoretical models are extralinguistic entities that represent target systems. They can be understood either as specific mathematical structures, for example state-spaces and transitions between states, as proposed by van Fraassen and Suppe, or, following the Tarskian tradition previously mentioned, as set-theoretical structures that satisfy theoretical statements. The general idea of the semantic view is that the theoretical statements found in scientific textbooks do not constitute the theory, but merely describe, in a particular language, the family of theoretical structures that constitute it. As Suppe (1989, p. 4) explains, the semantic view “construes theories as what their formulations refer to when the formulations are given a (formal) semantic interpretation”. There can be more than one way of describing these models, and more than one formulation of a theory.

Note that there is a straightforward tension between the semantic view, according to which scientific theories are *not* linguistic entities, and the idea that scientific theories are capable of being true or false, which is entertained by scientific realists and by some empiricists (including van Fraassen). Models are not generally said to be true or false. Instead, they are said to be accurate or inaccurate, or good or bad. So theories cannot only be collections of models, as acknowledged by many proponents of the semantic view, but they must be qualified by “various hypotheses linking those models with systems in the real world” (Giere 1997, p. 85), or by “a theoretical hypothesis claiming that real-world phenomena [...] stand in some mapping relationship to the theory structure” (Suppe 1989) (see also van Fraassen 1985).

This link between models and real systems can be analysed in terms of similarity (Giere 1997), or, in a structuralist spirit, in terms of isomorphism or partial isomorphism between the structure of the model and the (extensional, causal or modal) structure of its target (da Costa and French 2003; Bueno and French 2011; van Fraassen 1980). Concerning the relation to experience in particular, the idea that is generally put forth is that particular models of the theory, or their “empirical

substructures”, are directly compared to data models that represent particular phenomena, and again, mathematical tools such as morphisms can be put to work to analyse this relation.

2.2.3 *The Moderate Semantic View*

The semantic view can be seen as a way of bypassing the semantic problems associated with linguistic approaches towards science, in particular the issues affecting correspondence rules that plagued the statement view. However, one might suspect that this boils down to restricting oneself to the austere language of set-theory, and that correspondence issues are merely relegated to the philosophy of experimentation, having to do with the way target systems are identified and the way data models are extracted from experiments and compared to theoretical models.

An important criticism of purely structuralist approaches is that they are under-specified: many objects are similar or isomorphic to many other objects in various ways, so claiming that models are similar to what they represent is not saying much. As many commentators have observed, it is reasonable to assume that language must play a mediation role at some point in order to fix the intended target of representation for models, and incidentally, the domain of application of theories. For example, someone assuming that the model–target relation is some kind of isomorphism should provide a set of “important” objects, properties and relations, the structure of which is preserved by isomorphism, to avoid triviality (Ainsworth 2009). As for similarity, it can be argued that making explicit in which respects two objects are similar also requires linguistic mediation (Chakravartty 2001). In general, language is needed to say what the model is about (Thomson-Jones 2012; Frigg 2006; French and Saatsi 2006).

One can also presume that theoretical language must be interpreted consistently across various uses of the theory, which entails another important aspect: theories are more than mere collections of disparate structures; these structures are organised by a common vocabulary. It is important that what is called “electron” in one model correspond, in a way or another, to what is called “electron” in another model. As suggested by Halvorson (2012), theories might be better understood as topologies rather than collections of models.

So, we can say that a theory is a family of models organised by a common vocabulary. This much does not necessarily contradict the spirit of the semantic view. It leads to what we might call a moderate semantic view. In particular, it leads one to assume that a model is not a bare abstract structure, but that it comprises a certain connection with reality that is secured by linguistic reference. Assuming this moderate stance, the difference between the statement view and the semantic view could seem overstated, since there is not much difference between providing a set of statements and providing the set of structures satisfying these statements (but see van Fraassen 1989, p. 211). Yet, the focus on models and structures is still

present, and it induces a different philosophical approach towards science which has arguably proved fruitful.

According to its defenders, this way of construing theories offers a finer analysis of theory equivalence, of inter-theory relations and of the relation between theory and experience, by means of mathematical tools, than the statement view. It allows, for example, these analyses to operate at the local level of particular models rather than at the level of whole theories.¹ This could be compared to the way translating languages generally operates at the level of particular sentences rather than at the level of whole languages.

The focus on models has also proved fruitful in a way that was not necessarily intended by early defenders of the semantic view. It has allowed for the introduction of pragmatic considerations in philosophical accounts of the functioning of science. These considerations are the focus of the next section.

2.3 From Semantics to Pragmatics

This book rests on a conception of scientific theories that is similar to the semantic view, in that it is primarily focused on models. However, there is more than one way of developing a model-based conception of scientific theories, and early proposals within the semantic view did not remain unchallenged. Various authors sharing with the semantic view its emphasis on models as the main representational units in science started to defend more pragmatist-oriented views towards the end of the twentieth century (perhaps starting from Cartwright 1983) (see Winther 2015, section 4). The main characteristic of these approaches is an emphasis on scientific practice and on the contextual use of theories and models, rather than on formal aspects and abstract relations between theories and the world. This shift is sometimes achieved at the cost of giving up on providing a purely formal framework of science.

Note that many of these pragmatic aspects are endorsed by proponents of the semantic view, notably Giere and van Fraassen, and it is not obvious that there is a sharp distinction between two incompatible approaches here. Following Suárez and Pero (2019), we could talk about a representationalist semantic view, as opposed to a structuralist semantic view.

Let us present these pragmatic aspects. It can be helpful to distinguish three levels of analysis: the model–theory relationship, the model–world relationship and the model–experience relationship.

¹Rosaler (2015) provides such a local analysis of inter-theory reduction.

2.3.1 *The Model–Theory Relationship*

Concerning the relation between theories and models, one of the main observations is that models enjoy a certain autonomy from scientific theories (Morgan and Morrison 1999). The construction of a model based on a theory is not a systematic procedure. It often involves incorporating domain-specific assumptions, phenomenological laws or particular values for dynamical constants that are derived from experimental data. Furthermore, the models constructed from a theory are not always strictly speaking models of that theory, in the sense of structures that would satisfy theoretical laws. This is because scientists often use approximations that distort theoretical laws, such as perturbation techniques in quantum mechanics, or simplifying assumptions that are incompatible with said laws, for example, that the sun is fixed in a referential frame in a Newtonian model of the solar system. Scientists also combine incompatible theories, for example when modelling a quantum system in a classical environment.

Idealisations can be accommodated by claiming that idealised models can always be “de-idealised”, which only makes them better (McMullin 1985). Scientists would use idealised models because they are easier to use, but these departures from the theory would not make much practical difference, and they could be motivated by the theory itself, or by empirical observations. However, Cartwright et al. (1995) argue on the basis of an examination of the London model of supraconduction that this strategy is not always applicable. A scientific model can be independent from the theory in the sense that it “was not built as a de-idealization of high-level theory by improvements legitimated by independently acceptable descriptions of the phenomena”, but that some of the postulates it incorporates are somehow ad-hoc (Suárez and Cartwright 2007). According to Cartwright et al., this model autonomy would suggest that theories are not really collections of models, but rather tools for model construction, which constitutes an important departure from the traditional understanding of the semantic view. In this view, our representation of the world is not necessarily unified by overarching theories.

Note, however, that Cartwright et al.’s analysis of the London model is controversial (see for example French and Ladyman 1997). A recent examination of this historical case by Potters (2019) shows that indeed, ad-hoc hypotheses were involved in the construction of this model, but that they constituted a “programme” that was open to interpretation, and that a precise theoretical account of these new hypotheses was called for before the model and its interpretation could be stabilised. So, it seems that theories are a bit more than mere “tools”, and that theoretical unification remains an important aim of science. Nevertheless, this case shows that interactions between theories and models in scientific practice are not unidirectional.

One way to deal with idealisations is to simply relax the model–theory relationship. According to Giere (1999), the models of a theory should not be identified with the set of all the structures satisfying the general laws of the theory in the sense of Tarski. For example, not all types of forces are admissible in Newtonian mechanics. Only a few of them, such as electric forces, gravitational forces or

frictions, are apt to represent particular phenomena. Giere proposes that the family of models of a theory is limited, but open-ended (new forces could be added), and that it is hierarchical.² Higher levels in the hierarchy correspond to less specific models that can be made more specific in various ways, by enriching the dynamics with new components, for example, going from a harmonic oscillator to a damped oscillator, or by specifying the value of dynamical parameters or initial or boundary conditions. What unifies models together, at any level in this hierarchy, is some kind of similarity rather than a formal relation, which leaves room for idealisations.

In this picture, any abstract model can be associated with a family of more concrete models, which correspond to the different possible ways of making the model more precise. By analogy, one could think of scientific theories as very abstract models, the highest level of abstraction perhaps. This gives a way of understanding in what sense they can be considered families of models: because these very abstract models can be made more specific in various ways. It is important to note that each way of selecting a path down the hierarchy is associated with a focus on particular types of phenomena in the world. In other words, this hierarchy is empirically informed.

This does justice to the idea that models are partly autonomous from theories, that they act as mediators between the theory and the world, and this leads us to a consideration of the model–world relationship: what does it take for a model, whether abstract or concrete, to represent something?

2.3.2 *The Model–World Relationship*

If models are the main representational unit in science, understanding scientific representation does not necessarily boil down to a problem of philosophy of language. Maybe scientific representation is closer to pictorial representation than it is to linguistic representation. However, more should be said about this. What is the relation between a model and the world in general, and between a model and experience in particular? This question is of crucial importance for developing an empiricist position.

As said earlier, defenders of the semantic view often analyse the model–world relationship in terms of similarity or in terms of various types of morphisms, possibly mediated by theoretical language. These accounts are naturalistic, in the sense that users of representations do not play a constitutive role in the representation relation. But naturalistic accounts have been criticised for various reasons.

Suárez (2003) argues that having similarity or isomorphism between a model and an object is neither necessary nor sufficient for a model to represent an object. It is not sufficient because it has the wrong logical properties (as observed by

²A hierarchical construal of theories has also been proposed by Sneed (1971).

Goodman 1968 with regard to representation in general): the representation relation is directional (irreflexive and asymmetric), while similarity is not. For example, the object represented by a picture is not said to represent the picture, even though it is similar to it. Similarity or isomorphism are not necessary either, because misrepresentation is possible: one should distinguish representation simpliciter and accurate representation. But if a model represents its target in virtue of being similar to it in all relevant respects, then there is no room for inaccuracy (but see Bueno and French 2011 for a response to these arguments). This points to the idea that users and their intentions play an important role in establishing the representation relation, ensuring the directionality of representation and allowing for misrepresentation (I will say more about user-centred accounts of representation below).

We have seen that claiming that a model and its targets are similar or isomorphic requires specifying in which respect they are similar or isomorphic. The problem of misrepresentation could in principle be solved by distinguishing two kinds of relevant aspects: the ones that count for assessing if the model represents its target, and the ones that count for assessing if the model is accurate. Perhaps, for example, a model of the solar system represents its target in virtue of representing the right number of massive bodies, and it is accurate in virtue of assigning the right positions to these massive bodies. But it is questionable whether there is an objective, a-contextual way of drawing the line between the properties that “matter for representation” and the properties that “matter for accuracy”, and this can be seen as a reason for incorporating contextual elements in our account of representation. After all, such a distinction, if it existed, would seem to rest on a difference in accessibility, since we would naturally expect a scientist to be capable of knowing that her model represents a target, even if she does not know whether it represents it accurately. This does not necessarily mean that one cannot be wrong about the fact that a model represents a target in particular, but that this fact should be assertible in principle. So at least the distinction would depend on our epistemic position.³

The possibility of misrepresentation is certainly an important criterion for a viable account of representation in the context of the debate on scientific realism, for realists and empiricists alike. If misrepresentation was not possible, no theory could fail to be empirically adequate about the phenomena it represents, nor could it fail to be true, because representing a phenomenon and representing it accurately would amount to the same thing. This would trivialise the positions.

The possibility of misrepresentation is also important because of the widespread use of idealisations in science. We have seen that models sometimes distort theoretical laws. They also often caricature their targets, and sometimes incorporate fictitious entities for explanatory purposes, which look like instances of deliberate misrepresentation. Examples of this are frictionless planes, infinite gases or point masses: even if not incompatible with the laws of the theory, these assumptions are known to be false of the target system.

³I have developed this point in Ruyant (2020).

Pragmatist philosophers of science generally put emphasis on the sensitivity to purposes of modelling practices, in particular in the context of idealisations (Bailer-Jones 2003; Mäki 2009; Giere 2010a). It can be argued, for instance, that the way scientists idealise targets of representation depends on the variables they are interested in, on the type of activity (explaining, predicting) and on the levels of precision that they are expecting from their models. Taking the example of false assumptions in economics, Mäki (2009) argues that idealisations and fictions are used by scientists to neutralise uninteresting factors, sometimes by construction. This means that the model–world relation will be relative to a given perspective on phenomena adopted by users of the representation—a theme developed by van Fraassen (2008) and Giere (2010b).

2.3.3 *The Model–Experience Relationship*

Idealisations are sometimes accompanied by suitable interventions, the aim of which is to *physically* neutralise uninteresting factors, for example, when creating vacuum in experiments of free fall. This leads us to a final pragmatic aspect: the interactive and practical nature of experimentation. This aspect does not concern theoretical representational activities, but in experimental activities, establishing a representation relation between a model and a concrete object often involves controlling the object and its environment and performing various manipulations. These processes can rest on tacit practical knowledge, a learnt ability to use a telescope or a microscope for example, rather than on explicit or formal theory–world relations. As observed by Hacking (1983), telescopes had been used for centuries before scientists had a complete theory of their functioning. Contextual inputs are involved in these processes: for example, knowledge of specific sources of noise in a laboratory, or instrument calibration.

Some of these aspects can be seen as a means for scientists to *eliminate* contextual variations in order to reach cross-contextual stability. Bogen and Woodward (1988) propose to distinguish between data and phenomena on this basis (a distinction similar to van Fraassen’s distinction between phenomena and appearances, but see van Fraassen 2008, p. 376, note 14). Experimental data are cleaned up and synthesised using statistical techniques before they are compared to theoretical models. Modelling specific sources of noise is not required for these processes to be considered reliable. According to Bogen and Woodward, these processes are a means of accessing the phenomenon that causes the data, and the role of a model is not to account for messy data, but for phenomena. This means that the relation between a model and the world is not necessarily a contextual relation, in so far as the model would not represent brute data, but rather cross-contextually stable,

reproducible phenomena, accessible by means of different kinds of instruments.⁴ Feest (2011) further distinguishes between *surface phenomena* (or reproducible data patterns) and *hidden phenomena* (regularities between different types of experiments). She argues that stabilisation concerns the fit between these two types of phenomena, by means of a mutual adjustment between our classifications of them, and that it involves an interplay of conceptual competences (an ability to identify surface phenomena) and validation techniques.

In these views, the purposes of users could be taken into account for identifying the relevant phenomena. However, they could be seen as selective rather than as constituting the represented phenomena. Nevertheless, an important practical aspect remains. Our empirical access to stable phenomena with the mediation of messy data is not entirely captured by formal theories that would represent an experimental context in all its details. It also rests on practical or conceptual abilities and on efforts from experimenters to stabilise these phenomena. One could entertain the idea that mature theories and models result from a mutual adjustment between theoretical and practical aspects (exactly in the way as, as we saw, the model–theory relationship is not unidirectional). Hacking (1983, ch. 9)’s analysis of the development of thermodynamics, where “practical invention [...] gradually leads to theoretical analysis”, points in this direction (see also Chang 2004).

It should also be noted that the direction of fit between models and the world is not necessarily from the model to the world. Sometimes, and in particular in applied science and engineering, the world must comply to the model rather than the other way around. This is true even in theoretical science, when a physical system is carefully prepared to correspond to a model that was designed to test interesting implications of a theory: for example, testing Bell’s inequalities with an implementation of the EPR thought experiment. This could suggest that a model is not a mere description of a user-independent state of affairs, but rather a guide for interacting with reality, and that interpreting a model must in general involve referring to specific manipulations on the represented object. To be clear, this does not necessarily mean that no formal account of the model–phenomena relation can be constructed, but an ineliminable reference to intentional or user-centred aspects could be involved in this account. This interventionist aspect is a traditional theme of the pragmatist tradition, from Dewey’s criticism of the “spectator conception of knowledge” to more contemporary authors (Hacking 1983; Chang 2014; Brown 2015).

One could claim, by analogy with speech-acts such as promises or orders, that scientific models are performative. I borrow the notion of performativity from the philosophy of language. The general idea is that a speaker’s intention, what the speaker *does* by uttering a sentence, is fully part of what is meant by the

⁴Note that Bogen and Woodward’s distinction between data and phenomena has generated controversy, in particular regarding the independence of phenomena from agents (McAllister 1997) or the distinction between data and phenomena itself (Glymour 2000) (see Woodward 2011 for a response).

utterance. Although a scientific model is not exactly analogous to an assertion, the fact that the direction of fit between the model and the world can be one way or the other could let us think that intentional aspects, or what the model purports to achieve, must be taken into account to understand what constitutes the representation relation, that models are not purely descriptive, and that user purposes are not merely selective, but constitutive. This move can be resisted: it could be argued that intentions are merely instrumental in bringing about the phenomena in which we are interested, and that once this is done, a standard formal account of the model–phenomena relationship can be given without reference to user-centred aspects (this controversial point will come up again in Chap. 4). But at least, the fact that representational activities are not easily captured by a simple correspondence between the model and a pre-existing user-independent state of affairs gives us *prima facie* good reasons to move away from purely naturalistic accounts, and the remarks above concerning the practical competences involved in experimentation reinforce this idea.

Although model-based approaches towards scientific theories purport to distance themselves from linguistic approaches, there is a strong parallel with philosophy of language. Consideration of contextual use is analogous to the focus on pragmatics initiated by philosophers of natural language in reaction to formal semantics. Pragmatics is interested in the way context affects meaning, in particular with implicatures, indexical aspects, modulation and performativity. Notably, this part of philosophy of language is motivated by the fact that meaning is sensitive to the intentions of speakers and to salient aspects of the context, and that the direction of fit is not always word-to-world. Many of these aspects have their analogue in the philosophy of science literature (see for example van Fraassen 2008's use of the notion of indexicality). The account of scientific representation that I will present in Chap. 3 draws on this analogy.

2.4 From Users to Communities

We have seen so far how philosophers of science have moved from a linguistic to a model-based understanding of theories, and how this has opened the way for a consideration of pragmatic aspects.

All the aspects presented above, the directionality of representation, the possibility of misrepresentation, the sensitivity to purposes and the practical aspects of experimentation, point to the idea that the user of representation plays a central role in establishing the representation relation. However, this focus on practice, although informative, seems to threaten a certain unity of science, starting with the very possibility of identifying scientific theories: if science were nothing but a patchwork of representational activities involving disparate models for particular purposes, the axiological debate of scientific realism would be undermined, in so far as science could not be characterised by one unified aim associated with the production of scientific theories.

Some philosophers are willing to embrace this conclusion, by adopting a “dappled world” view, and assume that science is indeed constituted of disparate representations of particular domains that are merely partially unified (Cartwright 1999; Braillard et al. 2011). They endorse pluralism. However, giving up on conceptual unification being an important aim of science would be unsatisfying to many philosophers. After all, great theoretical achievements can often be seen as resolving tensions between various domains of knowledge: for example, relativity theory resolving the tension between Newtonian mechanics and electromagnetism.

I am convinced that the pluralist conclusion can be avoided by providing a clear articulation between the contextual and the communal components of representation. I now turn to this theme. Let us start with a presentation of the user-centred accounts of scientific representation that have been proposed to account for the pragmatic features examined in the previous section.

2.4.1 User-Centred Accounts of Representation

The strategy that is commonly adopted to take into account the aspects mentioned in the previous section, notably concerning the possibility of misrepresentation, the directionality of representation and its sensitivity to purposes and practical aspects, is to assume that the user plays a role in representation: representation is a three-place relation between a user, a source of representation and a target of representation (and possibly other aspects, such as purposes and audiences). This strategy has been defended by some proponents of similarity and isomorphism (Giere 2010a; van Fraassen 2008). So the question is now: what does it mean for a user to represent something using a model?

Some user-centred accounts of scientific representation are deflationary in spirit. According to Hughes (1997), representation in science involves three aspects: (i) a denotation of the target by the model, (ii) demonstrations performed on the model by its users, and (iii) an interpretation of the results of such demonstrations in terms of the target. According to Suárez (2004, p. 773), “A represents B only if (i) the representational force of A points towards B, and (ii) A allows competent and informed agents to draw specific inferences regarding B”. Representational force is “the capacity of a source to lead a competent and informed user to a consideration of the target” (p. 768), which includes having a denotational function (Suárez 2015, p. 44). These accounts concern scientific representation in general. Nothing particular is required concerning the nature of the vehicle or its relation to the target: it does not need to be an abstract structure for instance, nor to be similar to the target in any respect. These minimalist conceptions account for the directionality of representation, as well as for the possibility of misrepresentation: a representation is accurate if the inferences it allows lead to true conclusions, which need not be the case.

According to Suárez, although these conditions are not necessarily sufficient, because other norms could be involved in specific communities, nothing more can be said about representation in full generality.

Contessa (2007)'s account is also inferentialist, but purports to be more substantial. It is cast in terms of interpretation. Interpreting a model means, for a user, taking specific properties, relations or functions of the model to denote properties, relations and functions of the target. In brief, this means providing a mapping between relevant parts of the model and target. This, according to Contessa, is enough to fulfil Suárez's conditions of representation, because it allows the user to make inferences on the target (such as: if *a* is in the extension of *b* in the model, then the object denoted by *a* has the property denoted by *b*). Contessa claims that no more conditions are required. In particular, he assumes that any model–target mapping will do: to borrow his example, Rutherford's model of the atom could be used to represent a hockey puck sliding on ice, taking the electron to denote the puck and the nucleus to denote the ice. Although inaccurate, the model would count as a representation of the puck. This account is therefore more substantive, but also more liberal than Suárez's account.

An account of scientific representation that is particularly minimalist and liberal is provided by Callender and Cohen (2006). Drawing inspiration from Grice's reduction of linguistic meaning to mental states, they argue that scientific representation is based on stipulation only. A model represents its target because its users take it to represent this target: "representation is constituted in terms of stipulation (plus an underlying account of representation for the mental states subserving stipulation)" (p.77, footnote 7). According to them, all the particular characteristics of vehicles of representation in science, perhaps their similarity with their targets for instance, would not constitute the representation relation between the vehicle and the target. They would only correspond to features that are generally considered desirable for the purposes of the scientific community. They claim that a lot of debates that were framed as debates on the constitution of scientific representation are actually debates about pragmatic or demarcation issues.

Finally, some authors have recently proposed fictionalist accounts of representation. According to Godfrey-Smith (2007, p. 735), models are "hypothetical concrete things". Thus, these hypothetical objects can have properties that are normally associated with actual concrete objects, such as spatio-temporal properties, even when no concrete object is actually represented (typically, *the* model of the pendulum does not represent one particular pendulum), and this explains why scientists talk about such systems *as if* they were real. According to Frigg (2010), who takes inspiration from Walton's pretence approach to fictions in the philosophy of art, models are props in games of make-believe: they are constituted of instructions that prescribe us to imagine certain things using principles of generation. According to Toon (2010), these instructions for imagination directly concern actual objects, not imaginary ones. Similarly, Levy (2015, p. 741) proposes that models provide an "imaginative description of real things", and these descriptions can be partially true. This account purports to solve the difficulty in making sense of a comparison between actual and

fictional entities. Fictionalist accounts share with Callender and Cohen the idea that representation reduces to the mental states of the user.

These accounts of scientific representation, with their focus on users, are certainly able to accommodate the pragmatic aspects of the model–world and model–experience relations examined in the previous section. However, they could seem unsatisfactory in some respects. Arguably, naturalistic accounts of scientific representation are motivated by the intuition that the representation relation is not an arbitrary matter, so that in principle, something substantial can be said about the relation between sources and targets of representation. Although these naturalistic accounts present their difficulties, the intuition remains, and it can serve as a basis for criticising the deflationary accounts that have just been presented.

2.4.2 *Anything Goes?*

Callender and Cohen’s approach has been criticised by many authors (Toon 2010; Bueno and French 2011; Peschard 2011; Ducheyne 2012; Liu 2015; Frigg and Nguyen 2017; Boesch 2017). One type of criticism, notably by Liu, focuses on the fact that scientific models have an epistemic function. They “allow their users to have access, in however simplified or specialized manners, to aspects of their targets, which fulfil, in a broad sense, an epistemic role” (Liu 2015, p. 47). This makes epistemic representation in general, and scientific representation in particular, markedly distinct from symbolic representation: the function of a symbol (for example the logo of an airline) is merely to pick out its referent, which can be achieved by conventional stipulation. Any resemblance between a symbol and its target is merely pragmatic. With epistemic representations (for example, city maps), the vehicle–target connection is not entirely conventional, because this connection is essential to the realisation of the function of the vehicle. Liu remarks that epistemic vehicles generally contain symbols that help secure the relation to the target. But these symbols are carefully selected and organised to serve an epistemic function, whereas basic symbols do not need to contain other symbols.

Another type of criticism that will be important for our purpose concerns the communal aspect of representation. According to Boesch (2017), what a scientific model represents cannot be fixed by a mere stipulation by any particular scientist, because it depends on licensing by the community. Taking the example of the Lotka–Volterra model, Boesch argues that in order to know what a model represents, one has to examine its history: how and why it was constructed, how it was received by the scientific community and how it is currently used. This examination reveals that mere stipulation is not sufficient for a model to be licensed as a representation of something else, but that “the construction of the model [...] has been responsive to certain theoretical and empirical aims” (p. 976). According to Boesch, this licensing aspect is not pragmatic, but constitutive: a model is a scientific representation of its target because it has been carefully constructed to function as such, taking into account empirical and theoretical aims.

Similar kinds of criticisms can be addressed to Contessa's account in terms of interpretation. Although this account makes representation distinctively epistemic, in that the function of the model is to allow inferences, it does not rest on much more than stipulation, since any interpretation will do. This liberal stance has been criticised by Bolinska (2013), who claims that more is required for a vehicle to represent a target system. In particular, the agent must aim to faithfully represent the target system using that vehicle, and the vehicle must be informative about the target. An arbitrary mapping does not satisfy these conditions.

In a similar vein, fictionalist accounts have been criticised for not explaining how models produce knowledge about their targets. More than pretence seems to be involved in representation, and fictionalist accounts do not give us a means of discerning between the propositions that are held true or false by modellers (Poznic 2016). Knuuttila (2017) argues that fictionalist accounts lead to a problem of coordination between the imagination of various scientists. In response, one could focus on intersubjective principles of generation, but then it becomes unclear to what extent imagined objects have an added epistemic value (furthermore, many "principles of generation" go beyond our capacities for imagination, and require the use of computers Weisberg 2013).

What is called into question by these various criticisms, Bolinska's remarks on Contessa's account, Knuuttila and Poznic's criticisms of fictionalist accounts and Liu and Boesch's remarks on Callender and Cohen, is the viability of an account of representation that would be entirely based on the attitudes and purposes of the user. It is not true that anything goes: there are external constraints on what counts as a representation of something.

This aspect is certainly important for the purpose of this book, which is to provide and defend an empiricist account of scientific activity, in particular its aims and possible achievements. Paying attention to local experimental practices is, of course, indispensable for an empiricist. However, the axiological debate on scientific realism is not concerned with the purposes of particular scientists, but with the purposes of science in general. This means that the debate should be primarily focused on communal norms of representation, assuming that these norms are built around common values that bring together and animate the members of the scientific community. These norms have the capacity to account for the unifying power of scientific theories in a way that is compatible with a pragmatist stance. A good empiricist account of scientific representation should therefore explicate the articulation between contextual uses and communal norms. Let us now present some elements for this purpose, and outline the account of scientific representation that will be used in this book. This account will be fully developed in the next chapter.

2.4.3 Towards a Two-Stage Account of Representation

It is noteworthy that Callender and Cohen claim to draw inspiration from Grice. I have argued in Ruyant (2021) that Grice (1989, ch. 18)'s conception of linguistic

meaning has all the resources to answer the kind of worries that have just been raised concerning the distinction between symbolic and epistemic representation on the one hand and the communal aspect of representation on the other, and that Grice's analysis can be easily transposed to the case of scientific representation.

First, Grice's reduction of meaning to mental states is not as trivial as Callender and Cohen's reduction of representation to stipulation, so it might be possible to explicate epistemic representation in terms of mental states in a way that maintains the distinction between epistemic and symbolic representation. Secondly, Grice distinguishes utterance (or speaker) meaning and expression and word meaning. While the former reduces to the mental states of the speaker,⁵ the latter have to do with social values. Grice's idea is roughly to understand expression and word meaning in terms of speaker meaning by the mediation of communal values, and more precisely, in terms of appropriateness of use: for example, an expression means something in general if it is appropriate, in a community of speakers, to utter it to speaker-mean the corresponding belief.

I have suggested to transpose this analysis of meaning to scientific representation, and to distinguish two senses of epistemic representation:

- A concrete vehicle is used as an epistemic representation of its target in a given context if the user of the vehicle assumes or pretends that the vehicle is reliable for achieving a variety of purposes concerning the target (such as learning about the target) using interpretation rules.
- An abstract symbolic structure with general interpretation rules is an epistemic representation of a target or type of target if it is considered appropriate, within an epistemic community, to use a vehicle exemplifying or describing this abstract structure as an epistemic representation of this target or a particular instance of this type of target.

The first sense concerns contextual, user-centred representation, and it is quite liberal, but it is already distinct from symbolic representation, since it has an epistemic function. The second sense encapsulates communal "licensing" aspects, and brings more constraints on what counts as a representation of what. For example, if a user assumes that by using mathematical deduction rules on a set of equations on paper, and by interpreting symbols in these equations as referring to planets in the solar system, she can get to know the position of these planets at any time, then this set of equations *is used* as an epistemic representation of the solar system by this user. If, furthermore, such use is licensed by the community of scientists, then the abstract model used by this particular user by means of these equations *is* an epistemic representation of the solar system.

This conception accounts for the directionality of representation and for the possibility of misrepresentation at both levels. It can incorporate practical, contextual aspects, such as a sensitivity to purposes, in the first stage, but communal aspects are

⁵It reduces, more precisely, to an overt intention to induce a belief in an audience, and an intention that this overt intention be recognised as a reason to accept the belief.

involved in the second stage, and the way the two combine is explicit. This account is quite deflationary in spirit, but in contrast with other deflationary accounts, it has the merit of distinguishing explicitly the contextual and the communal levels. It also provides a clear articulation between concrete vehicles and abstract structures, each being associated with a different sense of representation, thus addressing another debate that was not mentioned in this chapter, concerning the ontology of models. More importantly, understanding the function of scientific models and theories in terms of norms has the capacity to account for the unity of science, conceived of as an activity directed towards a particular aim, without neglecting the variety of local practices. The aim of science would be to provide unified norms of representation that ensure success in contextual uses, whatever one's particular purpose. Incidentally, this understanding might restore the possibility of providing a formal framework for science, without neglecting contextual aspects.

In the next chapter, I present this two-stage account of epistemic representation in more detail, since this is the one that I will use in the rest of this book. Its main implication for our purpose is that the function of scientific theories and theoretical models should be construed in terms of norms constraining their particular applications, and the debate on scientific realism can be understood as bearing on the status of these norms and what their success implies.

2.5 A Tension Between Contextuality and Unity

In this chapter, I have presented various discussions, from which one can extract what I take to be the three main desiderata for a good account of scientific theories and representation: (i) it should be model-based, (ii) it should take into account pragmatic, practical and user-centred aspects (model autonomy, sensitivity to purposes, interactivity of experimentation) and (iii) it should take into account the communal nature of representation.

There is a certain tension between desiderata (ii) and (iii), which stems from the interplay of two important features of scientific activity. On the one hand, the central importance of experimental confrontation in science calls for a consideration of contextual representational uses; on the other hand, scientific theories and explanations have the capacity to unify disparate types of phenomena in a more or less systematic way, and this feature calls for a general, a-contextual treatment. Focusing too much on formal aspects, one might lose sight of the connection between theories and experience, but focusing too much on messy contextual practices, one might be unable to account for the unifying capacity of theories. However, understanding the communal representational status of scientific models and theories in normative terms can help overcome this tension.

In the context of the debate on scientific realism, note that empirical adequacy and truth, which are taken to be the aim of science by respectively empiricists and realists, are ideal, ampliative notions. By this, I mean that they are not restricted to actual representational uses: a theory that is ideally empirically adequate should

account for observable phenomena even when they are not actually observed by anyone. This is explicit in van Fraassen's definition of empirical adequacy, as we will see in Chap. 4. In principle, an ideally good theory should account for the free fall of a stone on a distant planet where no one will ever go, or it is not empirically adequate, and the same goes for truth. In this respect, conceptions of representation that are only applicable when a user is present and stipulates or denotes a target system are seriously limited. A good account of scientific representation must consider particular uses, but also take into account norms of representation that somehow transcend particular uses, and constrain even non-actual uses, so as to account for the axiological unity of science.

I believe that the right way of articulating unifying and contextual aspects is by relying on shared values and norms constraining particular uses rather than assuming a transcendental theory–world relation from the start, as the realist often does. The axiological debate can then be understood as bearing on these norms and their finality. This is why the articulation between communal norms and contextual uses provided by the two-stage account of representation sketched at the end of this chapter plays a crucial role in this work. In the following chapter, I will present this account in detail, and I will explain how the axiological debate between scientific realism and empiricism can be understood in this framework.

References

- Ainsworth, P. (2009). Newman's objection. *British Journal for the Philosophy of Science*, 60(1), 135–171.
- Bailer-Jones, D. (2003). When scientific models represent. *International Studies in the Philosophy of Science*, 17(1), 59–74. <https://doi.org/10.1080/026985903005238>
- Bailer-Jones, D. (2013). *Scientific models in philosophy of science* (vol. 43). Pittsburgh, PA, USA: University of Pittsburgh Press.
- Boesch, B. (2017). There is a special problem of scientific representation. *Philosophy of Science*, 84(5), 970–981.
- Bogen, J., & Woodward, J. (1988). Saving the phenomena. *Philosophical Review*, 97(3), 303–352.
- Bolinska, A. (2013). Epistemic representation, informativeness and the aim of faithful representation. *Synthese*, 190(2), 219–234. <https://doi.org/10.1007/s11229-012-0143-6>
- Braillard, P. A., Guay, A., Imbert, C., & Pradeu, T. (2011). Une objectivité kaléidoscopique : Construire l'image scientifique du monde. *Philosophie*, 3(110), 46–71.
- Brown, M. (2015). The functional complexity of scientific evidence: The functional complexity of scientific evidence. *Metaphilosophy*, 46(1), 65–83. <https://doi.org/10.1111/meta.12123>
- Bueno, O., & French, S. (2011). How theories represent. *British Journal for the Philosophy of Science*, 62(4), 857–894.
- Callender, C., & Cohen, J. (2006). There is no special problem about scientific representation. *Theoria: Revista de Teoría, Historia y Fundamentos de la Ciencia*, 21(1), 67–85.
- Carnap, R. (1966). *Philosophical foundations of physics*. New York, NY, USA. Basic Books.
- Cartwright, N. (1983). *How the laws of physics lie* (vol. 34). Oxford: Oxford University Press.
- Cartwright, N. (1999). *The dappled world: A study of the boundaries of science* (vol. 36). Cambridge: Cambridge University Press.

- Cartwright, N., Shomar, T., & Suárez, M. (1995). The tool box of science: Tools for the building of models with a superconductivity example. *Poznan Studies in the Philosophy of the Sciences and the Humanities*, 44, 137–149.
- Chakravartty, A. (2001). The semantic or model-theoretic view of theories and scientific realism. *Synthese*, 127(3), 325–345.
- Chang, H. (2004). *Inventing temperature: Measurement and scientific progress*. Oxford: Oxford University Press.
- Chang, H. (2014). Is water H₂O? *Evidence, realism and pluralism*. No. 293 in *Boston studies in the philosophy of science*. Cambridge, England; New York: Springer. oCLC: ocn870412305.
- Contessa, G. (2007). Scientific representation, interpretation, and surrogative reasoning. *Philosophy of Science*, 74(1), 48–68.
- da Costa, N., & French, S. (2003). *Science and partial truth: a unitary approach to models and scientific reasoning*. Oxford Studies in Philosophy of Science. Oxford: Oxford University Press.
- Ducheyne, S. (2012). Scientific representations as limiting cases. *Erkenntnis*, 76(1), 73–89. <https://doi.org/10.1007/s10670-011-9309-8>
- Feest, U. (2011). What exactly is stabilized when phenomena are stabilized? *Synthese*, 182(1), 57–71. <https://doi.org/10.1007/s11229-009-9616-7>
- French, S., & Ladyman, J. (1997). Superconductivity and structures: Revisiting the London account. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 28(3), 363–393. [https://doi.org/10.1016/S1355-2198\(97\)00013-0](https://doi.org/10.1016/S1355-2198(97)00013-0)
- French, S., & Saatsi, J. (2006). Realism about structure: The semantic view and nonlinguistic representations. *Philosophy of Science*, 73(5), 548–559.
- Frigg, R. (2006). Scientific representation and the semantic view of theories. *Theoria*, 21(1), 49–65.
- Frigg, R. (2010). Models and fiction. *Synthese*, 172(2), 251–268.
- Frigg, R., & Nguyen, J. (2017). Scientific representation is representation-as. In H. K. Chao & J. Reiss (Eds.), *Philosophy of science in practice: Nancy cartwright and the nature of scientific reasoning* (pp. 149–179). New York, NY, USA: Springer International Publishing.
- Giere, R. (1997). *Explaining science: A cognitive approach* (4th ed.). Science and Its Conceptual Foundations Series. Chicago, LI, USA: The University of Chicago Press.
- Giere, R. (1999). *Science without laws*. Science and its Conceptual Foundations. Chicago, LI, USA: University of Chicago Press.
- Giere, R. (2010a). An agent-based conception of models and scientific representation. *Synthese*, 172(2), 269–281. <https://doi.org/10.1007/s11229-009-9506-z>
- Giere, R. (2010b). *Scientific perspectivism*. Paperback edn. Chicago, LI, USA: University of Chicago Press.
- Glymour, B. (2000). Data and phenomena: A distinction reconsidered. *Erkenntnis*, 52(1), 29–37.
- Godfrey-Smith, P. (2007). The strategy of model-based science. *Biology & Philosophy*, 21(5), 725–740. <https://doi.org/10.1007/s10539-006-9054-6>
- Goodman, N. (1968). *Languages of art*. Indianapolis, IN, USA: Bobbs-Merrill.
- Grice, P. (1989). *Studies in the way of words*. Cambridge, MA, USA: Harvard University Press.
- Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science* (vol. 22). Cambridge: Cambridge University Press.
- Halvorson, H. (2012). What scientific theories could not be. *Philosophy of Science*, 79(2), 183–206.
- Hughes, R. (1997). Models and representation. *Philosophy of Science*, 64, S325–S336. <https://doi.org/10.1086/392611>
- Knuuttila, T. (2017). Imagination extended and embedded: Artifactual versus fictional accounts of models. *Synthese*. <https://doi.org/10.1007/s11229-017-1545-2>
- Ladyman, J., & Ross, D. (2007). *Every thing must go: Metaphysics naturalized*. Oxford: Oxford University Press.
- Levy, A. (2015). Modeling without models. *Philosophical Studies*, 172(3), 781–798. <https://doi.org/10.1007/s11098-014-0333-9>

- Liu, C. (2015). Re-inflating the conception of scientific representation. *International Studies in the Philosophy of Science*, 29(1), 41–59. <https://doi.org/10.1080/02698595.2014.979671>
- Lutz, S. (2017). What was the syntax-semantics debate in the philosophy of science about? *Philosophy and Phenomenological Research*, 95(2), 319–352. <https://doi.org/10.1111/phpr.12221>
- McAllister, J. W. (1997). Phenomena and patterns in data sets. *Erkenntnis*, 47(2), 217–228.
- McMullin, E. (1985). Galilean idealization. *Studies in History and Philosophy of Science Part A*, 16(3), 247–273. [https://doi.org/10.1016/0039-3681\(85\)90003-2](https://doi.org/10.1016/0039-3681(85)90003-2)
- Morgan, M., & Morrison, M. (1999). *Models as mediators: Perspectives on natural and social science*. Cambridge: Cambridge University Press.
- Mäki, U. (2009). MISSING the world. Models as isolations and credible surrogate systems. *Erkenntnis*, 70(1), 29–43. <https://doi.org/10.1007/s10670-008-9135-9>
- Peschard, I. (2011). Making sense of modeling: Beyond representation. *European Journal for Philosophy of Science*, 1(3), 335–352. <https://doi.org/10.1007/s13194-011-0032-8>
- Potters, J. (2019). Stabilization of phenomenon and meaning: On the London & London episode as a historical case in philosophy of science. *European Journal for Philosophy of Science*, 9(2), 23. <https://doi.org/10.1007/s13194-019-0247-7>
- Poznic, M. (2016). Make-believe and model-based representation in science: The epistemology of Frigg's and Toon's fictionalist views of modeling. *Teorema: International Journal of Philosophy*, 35(3), 201–218.
- Rosaler, J. (2015). Local reduction in physics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 50, 54–69.
- Ruyant, Q. (2020). Semantic realism in the semantic conception of theories. *Synthese*. <https://doi.org/10.1007/s11229-020-02557-8>
- Ruyant, Q. (2021). True Griceanism: Filling the Gaps in Callender and Cohen's Account of Scientific Representation. *Philosophy of Science* (forthcoming). <https://doi.org/10.1086/712882>
- Sneed, J. D. (1971). *The logical structure of mathematical physics*. Netherlands, Dordrecht: Springer. <https://doi.org/10.1007/978-94-010-3066-3>
- Stegmüller, W. (1979). *The structuralist view of theories: A possible analogue of the bourbaki programme in physical science*. Berlin: Springer.
- Suppe, F. (1989). *The semantic conception of theories and scientific realism*. Chicago, LI, USA: University of Illinois Press.
- Suppes, P. (1960). A comparison of the meaning and uses of models in mathematics and the empirical sciences. *Synthese*, 12(2–3), 287–301.
- Suárez, M. (2003). Scientific representation: Against similarity and isomorphism. *International Studies in the Philosophy of Science*, 17(3), 225–244.
- Suárez, M. (2004). An inferential conception of scientific representation. *Philosophy of Science*, 71(5), 767–779.
- Suárez, M. (2015). Deflationary representation, inference, and practice. *Studies in History and Philosophy of Science Part A*, 49, 36–47. <https://doi.org/10.1016/j.shpsa.2014.11.001>
- Suárez, M., & Cartwright, N. (2007). Theories: Tools versus models. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 39(1), 62–81.
- Suárez, M., & Pero, F. (2019). The representational semantic conception. *Philosophy of Science*, 86(2), 344–365. <https://doi.org/10.1086/702029>
- Thomson-Jones, M. (2012). Modeling without mathematics. *Philosophy of Science*, 79(5), 761–772.
- Toon, A. (2010). Models as make-believe. In R. Frigg, M. Hunter (Eds.) *Beyond mimesis and convention: Representation in art and science, boston studies in philosophy of science*. Berlin: Springer.
- van Fraassen, B. (1980). *The scientific image*. Oxford: Oxford University Press.
- van Fraassen, B. (1985). On the question of identification of a scientific theory (a reply to “Van Fraassen's concept of empirical theory” by Pérez Ransanz). *Critica*, 17(51), 21–29.

- van Fraassen, B. (1989). *Laws and symmetry* (vol. 102). Oxford: Oxford University Press.
- van Fraassen, B. (2008). *Scientific representation: Paradoxes of perspective* (vol. 70). Oxford: Oxford University Press.
- Weisberg, M. (2013). *Simulation and similarity: Using models to understand the world*. Oxford: Oxford University Press.
- Winther, R. G. (2015). In: E. Zalta (Ed.) *The structure of scientific theories*. Stanford Encyclopedia of Philosophy, Metaphysics Research Lab, Stanford University.
- Woodward, J. F. (2011). Data and phenomena: a restatement and defense. *Synthese*, 182(1), 165–179.

Chapter 3

Contextual Use and Communal Norms



Abstract This chapter presents the conception of scientific representation and theories used in this book. Its aim is to clearly articulate the contextual aspects of representational uses and the unificatory power of scientific theories. This is done by means of a two-stage account of representation. The first stage involves user attitudes and purposes, which are formalised through the notions of context and interpretation, and introduces a notion of accuracy that captures experimental success. The second stage is focused on communal norms concerning particular uses. It is captured by the notions of general model and relevance. A scientific theory is understood as a family of general models. This account provides a good basis for framing the axiological debate on scientific realism, as bearing on the values involved in the selection of communal norms of representation, and in particular empirical adequacy.

3.1 The Two-Stage Account of Epistemic Representation

We have seen in the previous chapter that the use of scientific models is characterised by a sensitivity to purposes and by interventionist aspects. However, theories also have the capacity to unify various types of phenomena. There is a certain tension between these two characteristics of science. A good account of theories and representation should be able to combine them. This is particularly important from an empiricist perspective, because empiricism considers the relation between theories and experience to be a central component of a proper understanding of the aim of science, and at the same time, empiricist positions, conceived of as axiological positions, rest on the idea that there is such a thing as a unified aim for science, and not only a patchwork of local purposes. The aim of this chapter is to develop a positive proposal in order to address this issue.

There are two possible attitudes that one could adopt when confronted with this tension between unifying and contextual aspects. The first one is what we could call a theory-first approach: it consists in assuming a universal semantics for theories, interpreted as unified, objective descriptions of their domain of application, and then attempt to account for contextual uses in these universal terms. Perhaps, for

example, our best theories of physics would have a model of the universe such that all particular uses of these theories “fit inside”. The other attitude is a practice-first approach: it consists in accounting for the unificatory power of theories in terms of particular uses rather than the other way around. The analogue of this second approach in philosophy of language is an understanding of meaning in terms of use rather than in terms of formal, a-contextual semantics. This approach naturally leads to a consideration of communal norms as playing a unificatory role, in place of transcendental theory–world relations.

The role played by users in scientific representation and the practical aspects of experimentation examined in the previous chapter play in favour of a practice-first approach. The previous chapter ended with the sketch of a two-stage account of scientific representation that fits into this approach: the unificatory power of theories is explained in terms of contextual uses rather than the other way around, by appealing to communal norms of appropriate use. The main feature of this two-stage account is that it distinguishes clearly between these two levels: the first stage of representation concerns contextual representational use, which depends on the attitudes and purposes of particular users, while the second stage concerns general representational status, which has to do with communal norms. The two are also clearly articulated, since general norms constrain contextual uses.

It is now time to present this account in more detail so as to provide a solid framework for the construction and defence of the empiricist position to which this book is dedicated. This is the purpose of this chapter. I will explain what, according to a modal empiricist, scientific models are, how they are confronted with experience and how theories unify them. Since the present work is not about scientific representation, I will not provide a detailed argumentation in support of this account (for its main motivations, see Ruyant 2021), but I hope that it will be clear to the reader that it is sufficiently flexible to account for the main features of scientific representation presented in Chap. 2. Besides, it is quite deflationary in spirit, and I think that no unreasonable commitments are required for its construction. A scientific realist will most certainly need a more substantial account of scientific representation to flesh out her position. I will sketch a realist account in Chap. 6 for the sake of the discussion. However, It seems to me that even someone with a more substantial conception of scientific representation could accept that this account correctly describes the surface features of representational activities.

Note that representational vehicles in science are numerous: they include diagrams, photographs, computer simulations, scale models and mathematical equations, among other entities. All these vehicles of representation have an epistemic function, in the sense that they can be used to learn about their target. Other vehicles outside of science, such as city maps, also have this function. The two-stage account developed in this chapter concerns epistemic representation in general.

Having said that, My interest in this work lies primarily with scientific representation, and with theoretical models in particular, so most examples and many comments will be about representation by means of theoretical models. The last section of this chapter is dedicated to considerations that are specific to science.

One of the main contributions of this presentation for the rest of this book is the introduction of two functions, *accuracy* and *relevance*, each concerning one stage of representation. These two functions will be used to define empirical adequacy, taken to be the aim of scientific theorising, in the next chapter. Other notions, such as context and interpretation, will be used throughout the book as well.

3.2 First Stage: Contextual Use

The first stage of the two-stage account of epistemic representation is concerned with the use of vehicles by agents in relation to concrete targets with which these agents interact for particular purposes. This can be, for example, the use of a map (marks on paper) to navigate a city. In science, this means experimental contexts, for example when testing the validity of a hypothesis or when developing technologies.

According to the two-stage account of representation, this first stage reduces to the mental states of the user: a concrete vehicle is used as an epistemic representation of a concrete target if and only if the user assumes, or pretends, that she can achieve any purpose within a finite range (such as going from one point to another in a city, or learning about some properties of a physical system) by means of the vehicle, using specific interpretation rules connecting the vehicle and the target.

This stage takes into account the sensitivity to particular purposes and the practical aspects of experimentation presented in Sect. 2.3. In this section, I provide semi-formalisations of these aspects by means of two notions: context and interpretation. Then I introduce a third notion, accuracy, which characterises experimental success.

3.2.1 Contexts

The notion of context will play an important role in this book. It is meant to capture the purposes that an agent could entertain with respect to an object, such as navigating between any two points in a city. A context specifies a target of representation, which is a concrete object or a set of objects that can be referred to by ostentation (for example, points of interest in a city, or a pendulum), and a set of properties of interest (connections between these points, the position of the pendulum along an axis). In what follows, I use property in a broad sense, which includes any n-ary relation.

In a scientific context, a user will often express her purposes in theory-laden terms: for example, she will be interested in knowing whether the spin-z of a particle is up or down relative to the reference frame of her laboratory. The fact that spin is not directly observable should not bother us. I assume that in particular contexts, theoretical properties are translated in terms of *appearances*, that is, in terms of

observations and manipulations, for example the result of a particular measurement, a flash on a scintillation screen, or perhaps a value in a data model that aggregates several results. This corresponds to the “local meanings” of these properties in these particular contexts.

These “local meanings” must presumably respect communal norms concerning the interpretation or operationalisation of theoretical vocabulary. The transcription of language into appearances can be fairly complex in the case of science. However, let us postpone the discussion of this aspect to Sect. 3.4.2. For now, we can simply assume that in science, a context is a theory-laden entity: the properties of interest that it specifies, although locally interpreted in terms of appearances, are expressed in a theoretical vocabulary. What is specifically contextual in this stage is not the way theoretical terms are operationalised, but the fact that users are interested in particular properties of the target and not in others.

These properties of interest will generally be coarse-grained properties, and sometimes aggregates. The user could, for example, be interested in the position of a pendulum with a precision of one centimetre, because of limitations of her measuring apparatus, or the user could be interested in the mean value of a quantity among a collection of objects. I assume that properties of interest are in principle accessible to the agent, and that they are discrete and in finite number.

These properties of interest can be classified into two categories:

- the properties that are fixed by practical and empirical inputs in context, and
- the properties that the user wants to inquire about.

The way properties of the first category are fixed can be understood by considering the case of maps, where inferences on which path to follow involve assuming a starting point and a destination point. The starting point, assuming it corresponds to where the user is located, is an empirical input given by the context, and the destination point is a practical input. Similarly, inferences afforded by a theoretical model can take empirical or practical inputs, such as a measured quantity corresponding to the initial state of a system or a desired result (in which case the inferences could concern what operation should be performed to get this result). One can also take the type of phenomena represented, that is, the specific physical configuration that the user is interested in (say, an electron in an electromagnetic field) to be fixed by the context in this sense. Just like other properties, it can be fixed either practically if this physical configuration is to be implemented by suitable manipulations, or empirically if the user is interested in an existing object. These inputs do not correspond to what the user wants to learn, but identify the object of representation, that is, the kind of entity about which the user wants to learn something. They are given by the context in the form of properties of interest of the first category.

The properties of the second category correspond to possibilities that are not necessarily actual. When using a map, a user wants to know which path leads more directly to her destination. Not all imaginable paths have this property, but it is a priori conceivable that they do. In the spin example given above, the user is interested in whether a particle has spin up or down, which will correspond to

two exclusive measurement outcomes in the context, and both outcomes are a priori conceivable. The context must specify these two possibilities in order to express what the user wants to know. An intuitive way of understanding this aspect is in terms of a closed question being asked, and this question can be expressed by specifying the list of all its possible answers: in our example, {spin-up, spin-down} is a way of expressing the question “what is the spin-z of the particle?”. A context is taken by the user to be compatible with a range of possible questions, and it specifies the possible combinations of answers to all these questions, assuming they could all be answered conjointly.¹

Note that the notion of possibility involved when talking about possible answers is purely epistemic, and has no metaphysical implications. It corresponds to what is conceivable given what the user knows about the target and given her discrimination abilities. If, for instance, the user plans to measure the position of a pendulum with a ten centimetres ruler with a precision of one centimetre, then there are exactly eleven conceivable outcomes, corresponding to the ten intervals that the ruler can discriminate plus the possibility that the position of the target is outside of this range. The user will want to learn from her epistemic representation which of these outcomes could occur, but all are a priori conceivable. Whether or not some of these possibilities are physically impossible is irrelevant at this stage. The notion of conceivability involved here is also distinct from the one entertained in metaphysics, for example when pondering, in philosophy of mind, whether a world identical to our world but without phenomenal consciousness is conceivable. This notion is distinct in two respects. First, we are here concerned with a limited set of properties of a particular, actual object, not possible worlds, and second, these properties must be empirically accessible in principle, which means that the conceivable properties specified by a context are connected to conceivable manipulations or observations on the target (not qualia or the like).

A context can be formalised by giving a set of mutually exclusive conceivable states in which the target system could be found, each conceivable state corresponding to a particular combination of coarse-grained properties and relations. In the case of dynamical systems, these properties will be indexed to time intervals, in which case we could talk about conceivable histories. In other words, the context *denotes* concrete objects and *describes* the conceivable properties they could have given a certain perspective on these objects. The properties of the first category above are present in all conceivable states, because the system must have them for empirical or practical reasons: they are epistemically or practically necessary. The properties of the second category vary from one conceivable state to another, all possible combinations being represented. I will assume this formalisation from now on.

¹In the case of quantum mechanics, there is a sense in which not all “questions” about a system can be answered conjointly, and we can take the context to specify a set of compatible “questions”, which corresponds to a perspective on the target adopted by the user. This could be formalised using Griffiths (2003)’s notion of framework in the consistent histories approach.

Definition 3.1 Context:= set of mutually exclusive conceivable coarse-grained states or histories for a concrete target system.

This formalisation can express what, in a context of use such as an experimental context, the user already knows and what the user could want to inquire about without presupposing what the results of inquiry would be, nor any constraint on these results. The use of coarse-grained or aggregate properties can express various levels of requirement, which can be useful to account for idealisations. In sum, the context captures a certain perspective on a target.

During an experiment, exactly one of the possible states or histories specified by the context will be realised. In science, this result will typically be represented by a data model. A context in science can therefore be understood as specifying a set of conceivable data models, one of which will be realised, which corresponds to what Suppes (1969) calls a model of the experiment.

3.2.2 *Interpretation*

In a context, the user manipulates a concrete vehicle, which can be, for example, a set of equations, to make inferences about the target system. The conclusions of these inferences can then be used to interact with the target in a certain way. The concrete vehicle used by the user describes or instantiates a symbolic structure (which can roughly be understood in Tarskian terms, that is, as a domain of objects in which the extensions of various theoretical symbols are specified). This is the case including for physical models, for example, the scale model of a molecule: the balls and sticks of this type of model have a symbolic function, and they instantiate a particular structure.

According to the two-stage account of representation, representational use reduces to the mental states of the user. It reduces, more precisely, to an assumption or pretence that the vehicle is a reliable means of achieving various purposes concerning the target by using systematic interpretation rules. These interpretation rules specify the denotation of symbols: for example, that x in a model of a pendulum denotes the position of a concrete pendulum along a particular axis, or more precisely, that the possible values that this variable can take denote the conceivable positions of the concrete pendulum along this axis. The symbolic structure described or instantiated by the vehicle constrains the inferences that can be made, and these inferences are about the entities of the real world denoted by the symbols. Purposes, such as learning about the target, are achieved by means of these inferences. Typically, in physics, a model incorporates dynamical constraints that determine the inferences that can be made concerning the various quantities that are denoted by particular symbols, for example, dynamical constraints on the possible trajectories of a pendulum. These inferences allows one to make predictions or to manipulate the target appropriately for various purposes.

In a first approach, this representation relation can be roughly formalised using Contessa (2007)’s notion of interpretation (see Sect. 2.4.1) as a mapping between objects, properties and functions of the model to objects, properties and functions of the target (see also Weisberg (2013, ch. 3.3)’s notion of construal). A user takes an axis of the coordinate system of the model to denote a physical direction in the laboratory, or the value of a variable to denote a measurable property the target could have (the position of the pendulum, the spin-z of an electron). Or the user takes particular symbols on a city map to denote metro stations.

Since denoted properties will be limited to the accessible properties of interest specified by the context, we can just as well say that the interpretation of the model is a mapping between components of the model and the target properties specified by the context: we can easily account for the denotation of objects and functions in terms of simple or complex properties. This mapping can be many-to-one in case the model is more finely grained than required by the context. It can be, for example, a mapping from every real numbers between 0 and 1 for a dimension in the model to the coarse-grained property of having a position measured in this interval for the target. Components of the model considered fictitious, not measurable or uninteresting can be excluded from the mapping.

The contextual interpretation of a vehicle is, formally speaking, a *projection* of the symbolic structure of the vehicle onto the context, and as said earlier, the context provides a perspective on an object. An analogy with projections on spatial perspectives can be helpful to understand what is going on, and the comparison can be quite literal in some cases. Imagine, for example, having a three-dimensional model of a pendulum, such that the pendulum can move in the three spatial directions. Imagine that you plan to measure its position along two axes only, by taking photographs of the pendulum from a single point of view at different times. In this case, the context is restricted to two variables for the target, associated with its position along the two relevant axes, which literally corresponds to a spatial perspective on the object. The interpretation of the model will map components of the three-dimensional model to these two position variables, which is just a projection of the three-dimensional model onto a two-dimensional perspective. In general, an agent will assume that a given model is applicable in more than one context, in the same way that a three-dimensional model can be projected onto more than one spatial perspective.

This account of the representation relation in terms of a mapping between components of a model and target properties is similar to Contessa’s. However, I wish to highlight some potential differences between the two, which directly stem from our understanding of contexts (but some of these aspects might already be implied by Contessa’s account):

Coarse-graining: The mapped properties can be coarse-grained properties or aggregates.

Modality:	The mapped properties can correspond to merely conceivable states for represented systems (epistemic possibilities); they need not be actual properties of the target. ²
Epistemic relativity:	Although expressed in a theoretical vocabulary, the properties denoted need not be natural properties: they can be relative to an epistemic perspective, and ultimately translated in terms that are relative to the user and her technical abilities (typically, the result of a measurement).
Performativity:	Incidentally, the interpretation can be performative: the denotation of a symbol can imply that a certain manipulation is performed on the system, or the symbol can directly denote a manipulation, such as preparing a particular state for the target.

Here is a toy example of a context and interpretation. Imagine you plan to throw a dice twice, and you are interested in predicting whether the results will be odd (*O*) or even (*E*). The context corresponds to all possibilities in this respect: $\{\{O, O\}, \{O, E\}, \{E, O\}, \{E, E\}\}$. You decide to use a model of the dice for making predictions. According to this model, the absolute difference between two successive numbers displayed by the dice is exactly 1. The model is described by equations such as $|r_{n+1} - r_n| = 1$ and the mathematical structure of the model is given by the full set of infinite sequences respecting this condition, including for example $\{1, 2, 3, 2, 3, 2, \dots\}$ or $\{4, 3, 2, 3, 4, 5, \dots\}$. Your interpretation of the model maps the first and second member of every sequence (r_0 and r_1) to the first and second results of the context, and it maps the properties of being a 1, a 3 or a 5 to the contextual property of being odd (*O*), and the properties of being a 2, a 4 or a 6 to the contextual property of being even (*E*). The model is “projected” onto the context, in order to fit your particular purposes. Assuming that you are confident that the model, interpreted in this way, is reliable for predicting the results of your actions with the dice, we can say that you are using this model to represent the situation.

I assume that this general account can be applied similarly to more realistic examples of representation, for example in scientific experimentation, even if the interpretation could be less trivial. As argued by Contessa, having an interpretation is enough for being able to make inferences about a target. The inferences concerning the target afforded by the model can be understood as the set of statements that the model satisfies, translated into statements about the properties of interest by means of the interpretation. In the illustration above, one can infer for example that if the first result is *O*, then the second result will be *E*. This is because the model satisfies the statement $r_0 \in \{1, 3, 5\} \rightarrow r_1 \in \{2, 4, 6\}$, which, given the interpretation, can be translated into the statement that if the first result is odd, then the second result is even.

²Remember, however, that this notion of conceivability has no metaphysical implications.

One could argue that this account of representation is too substantial, and that this notion of interpretation is a mere means, among others, of representing a target (see Sect. 2.4.1). Suárez (2004) notably puts emphasis on the plurality of representational means and vehicles: the latter need not be abstract structures, for instance. However, in this account, vehicles are not abstract structures. They describe or instantiate a structure, which eventually boils down to the fact that they constrain possible inferences. This account could be applied to the scale model of a bridge for instance, where components of the model would be mapped to properties of a real bridge.³ Furthermore, we are only concerned with the representation of concrete targets in this first stage, and not the representation of types or abstract entities. Something like denotation is needed to secure the relation with the target. It is hard to imagine how one could represent a concrete entity without the kind of denotation involved here, particularly given the inclusion of coarse-graining, modality, epistemic relativity and performativity.

3.2.3 Concrete and Abstract Interpreted Models

What I call a *concrete interpreted model* is the combination of a vehicle and its interpretation. Note, however, that many aspects of the vehicle associated with its physical constitution do not really matter for the representation relation, in so far as they do not affect the inferences that are afforded by the model. These are, at best, pragmatic features that facilitate inferences (for example, the colours and fonts used in a map, what Vorms 2011 calls a *format*), and at worse, irrelevant or undesirable aspects (stains on the map). For our purpose, it will be convenient to abstract away these features, and to focus only on the way the vehicle constrains inferences, that is, on the abstract symbolic structure that the vehicle exemplifies or describes.

This symbolic structure is relative to the interpretation: it is not the complete physical structure of the vehicle (whatever this could mean), but the structure exemplified by the component parts of the vehicle that are interpreted in terms of the target. In general, this structure will be modal or intensional: it will specify, for example, a set of possible combinations of fine-grained theoretical properties the target system could have at different times.⁴ In the illustration with a dice given above, the vehicle is a set of equations on a sheet of paper, and the symbolic structure it describes corresponds to the first two members of every sequences of results

³In the case of physical models, inferences can rest on physical processes, and not only on mental processes. In the case of computational models, they rest on calculations by a computer, which could be thought of as highly constrained physical processes. I am interested in theoretical models in this work, and I assume that only mental processes are involved in this case.

⁴In this sense, it is very close to a propositional content, since propositions can be analysed in terms of possible states of affairs.

respecting the conditions of the model.⁵ This symbolic structure is modal, because various possibilities are represented. Not all vehicles have this modal characteristic (arguably, maps do not), but some do. This is the case of probabilistic models, and of models of deterministic theories with free variables, for example when initial conditions are not specified.

We have seen that a context could be formalised as a set of conceivable coarse-grained states or histories for the system. What the structure of a model does, once interpreted in terms of the context, is bringing constraints on this set, so that particular possibilities are excluded. These are the possibilities that will not be mapped to any component of the model, because these possibilities do not appear in the model. In our illustration, the excluded possibilities are $\{E, E\}$ and $\{O, O\}$: these are the possibilities of the context that are not “permitted” by the model. To give another example, a context could specify that the position of a pendulum will be measured ten times in a row, and all possible sequences of outcome are conceivable, but the model will restrict these to the sequences that are “permitted” by its structure (the ones that respect the laws of physics for instance). Vehicles that do not have a modal aspect just correspond to the special case where only one possibility remains.

In the case of physical systems, all these ideas can be straightforwardly transcribed into talk of state-space, possible trajectories in state-space, dynamical constraints, initial and boundary conditions and partitions on state-space or histories for coarse-grained properties, using van Fraassen (1989, ch. 9)’s conception of a scientific model for instance (see Sect. 2.2.2). But the present account is more general.

What I call an *abstract interpreted model* is the combination of the modal symbolic structure exemplified or described by the vehicle and its interpretation in terms of the target of representation, that is, a structure, a context and a mapping between the symbols of this structure and the properties of interest specified by the context.

Definition 3.2 Abstract interpreted model := $\{S, C, I\}$ where S is a modal symbolic structure, C a context and I a mapping between S and C .

A model is accurate if the inferences it allows lead to true conclusions. Remember that these inferences correspond to the statements satisfied by the symbolic

⁵Only the first two members are included in this structure, because other members are not mapped to elements of the context. This is counterintuitive, because the model was initially presented in terms of infinite sequences. The same would go for a three-dimensional model of a pendulum in the case presented above: only a discrete, two-dimensional structure is effectively mapped to the context. One could integrate the full sequences or the full three-dimensional structure at this stage by considering that the user is confident that the model could be interpreted in a range of other conceivable contexts, or perhaps more neutrally, that other components of a model are potentially representational, and therefore fully part of the structure of the vehicle. This would not affect the following considerations. However, these aspects might be better accounted for at the communal level that will be addressed in the next section. After all, in the dice example, what is being represented in context is not an infinite sequence of throws.

structure of the model, translated into statements about the target by means of the interpretation. This means that conditions of accuracy are fully determined by the structure of the vehicle and by its interpretation in terms of the context, that is, by the abstract interpreted model that the vehicle instantiates. One could think of abstract interpreted models as giving the content, or “meaning” of concrete vehicles, by analogy with truth-conditional analysis of meaning in philosophy of language. This meaning is contextual, since the model is interpreted in terms of a context that depends on the particular purposes of the agent. Whether the conditions of accuracy of an interpreted model are met depends on the target. Let us now examine this aspect.

3.2.4 Accuracy

It is important, for the model to function as an epistemic representation, that the user assumes or pretends that the model is accurate. Pretence is sufficient, because a model can be tested, in which case there is no prior assumption of accuracy, yet inferences are made and operations are performed as if the model was a reliable guide. The user could also assume more than empirical accuracy. For example, if she is a realist, she could assume that the model, interpreted in a certain metaphysical way, correctly describes the fundamental nature of the target, or that the component of the models that are not actually mapped to the context actually correspond to hidden properties of the target. But this is not required for the representation relation to take place (otherwise, idealised models incorporating fictitious entities would not represent). In any case, these attitudes from the user are mere attitudes; they do not entail that the model is actually accurate, so misrepresentation is possible.

I assume that the notion of accuracy involved here is purely extensional. As noted earlier, scientific models sometimes have a modal or probabilistic structure, and epistemic vehicles bring constraints on possible states of affairs for the target system. For example, one could infer from a model that if the initial position of a pendulum were x_0 at t_0 , then it would be x_1 at t_1 , even if the antecedent is actually false. These possibilities are mapped to the context. I will defend in this work that such counterfactual reasoning is indeed warranted. But the notion of accuracy involved, what one could call modal accuracy, is a distinct notion from this one. We will need a careful argumentation to obtain it (including an adequate notion of possibility), which has to wait for the next chapters.

I am only concerned so far with a notion of accuracy that captures empirical successes, and empirical success can only be assessed with regard to actual properties. So, all statements of the form “if p then q ” can be read in the form of a material conditional and considered trivially true in cases where p is false.

With the present extensional notion, an interpreted model can be considered accurate if the actual coarse-grained state of the system, restricted to properties of interest, is among the states permitted by the model interpreted in terms of the context. For example, a model of a pendulum is accurate in a particular context if

the sequence of outcomes actually measured is one of the sequences “permitted” by the model. It is important to remember that a model is interpreted in terms of a context, and that therefore accuracy is relative, or perspectival, in two respects. It is relative to our epistemic abilities, because the properties specified by the context must be accessible, and it is relative to the purposes of the user, because the context specifies properties of interests (which include relevant degrees of precision). In our illustration, if after throwing the dice twice, we obtain an odd number followed by an even number, then the model is accurate, whatever these numbers are, because this result is permitted by the interpreted model. This is true even if the model would not be accurate in another context, for example if the user focused on the exact result of the dice instead of its parity. Finally, representational use is in general performative, and accuracy can be conditioned on the fact that certain operations are performed on the target, such as placing a pendulum in a particular initial state. It depends on our manipulations as much as on our observations.

In science, the actual state of the system will generally be represented by a data model, which will correspond to one of the possible states or histories specified by the context. The model is accurate if the actual state represented by the data model belongs to the set of conceivable states of the context that are mapped to components of the interpreted model. This comes close to van Fraassen (1980)’s analysis of empirical adequacy in terms of data models being “embedded” in the structure of theoretical models. In this case, the statements that are satisfied by the model, once translated in terms of the context, will be satisfied by the data model, which amounts to saying that the inferences afforded by the model are true of the target.

Taking stock, let us introduce an accuracy function that takes the interpreted model as a parameter, and gives a truth value. Note that only the abstract structure of the vehicle matters for accuracy, so this function will apply to abstract interpreted models.

Definition 3.3 *accuracy*(m): specifies whether the actual, coarse-grained state or history of the system represented by m is mapped (that is, permitted) by m .

The extensional aspect of this function could be a source of difficulty for interpreting probabilistic models. The status of probabilities is a thorny issue that will not be addressed in depth in this book. All I can do here is sketch a possible way of interpreting probabilistic models. A probabilistic model does not permit or exclude conceivable states, but it attributes a weight to them, so we can assume that accuracy is a matter of degree, and that an accuracy function assigns probabilities instead of truth-values to models. These probabilities would express the match between the frequencies of actual measurement results and the probabilities assigned to these results by models, taking into account the sample sizes (which roughly corresponds to p-values). The attitudes of the user could perhaps be analysed in terms of rational bets concerning these frequencies. As is well known, an identification of probabilities with frequencies is problematic. I will introduce a modal version of adequacy in the next chapter which could alleviate these worries. Whether or not this is enough for a pragmatist interpretation

of probabilities would require more analyses, but unfortunately, I will not address this issue in this work.

The notion of accuracy will play an important role in the next chapter, when the notion of empirical adequacy will be examined, since the accuracy of particular uses of a theory can be considered the ultimate benchmark for their acceptability.

3.2.5 *Model Choice*

We now have a full account of contextual representational use. It consists in a context, which specifies properties of interest, a concrete vehicle instantiating or describing a symbolic structure, and an interpretation of the vehicle in terms of the context, which is a mapping between the two. The context specifies a set of conceivable coarse-grained states for a target system, and the vehicle and interpretation restrict this set to a sub-set of permitted coarse-grained states. The interpreted model is accurate if the actual state of the target is among these permitted states. A vehicle is used as a representation of a target system if its user assumes or pretends that the vehicle thus interpreted is accurate.

At this first stage, not much is implied concerning how the user selects or constructs her model, how she interprets it, and why she assumes or pretends that it is reliable. It seems reasonable to assume that the selection and interpretation parts will involve an identification of the type of system that must be represented, and practical considerations (that the model should be easily computable for example), as well as background assumptions that are inherited from the epistemic community: how target systems of this type ought to be modelled, and so on. The user could rely, for this purpose, on a particular theory and associated stock models. That is, she can rely on communal norms at work in her community concerning the appropriate way of modelling particular targets. But this is not necessary for the representation relation to take place at this first stage. If an agent is superstitious and interprets patterns in coffee grounds in terms of lottery results, using interpretation rules just invented on this occasion, and if the agent assumes that this is a reliable means to win the lottery, then the coffee grounds constitute an epistemic representation of lottery results at this first stage.

In case the user relies on theoretical background knowledge for choosing or constructing her model, for example using Newtonian mechanics to predict the position of a pendulum, this knowledge (Newtonian mechanics) will not depend on the particular context, and it will remain unquestioned in the context. One could say that it plays the role of conceptual necessity for the user. But the communal norms that users rely on in particular contexts can in principle be questioned at a more general level, taking into account the successes or failures of particular uses, and they can possibly be revised. This is an important component of scientific activity. This leads us to the second stage of representation, which concerns the communal status of representational vehicles.

3.3 Second Stage: Communal Status

We now have an account of what it is for a user to represent something for epistemic purposes. The representation relation can be analysed in terms of an interpretation, that is, a mapping between components of a vehicle and relevant properties of a target. This is enough to establish a representation relation in the first stage of our account. However, this might not be sufficient to account for the representational status of epistemic vehicles, and in particular, of scientific models. What is missing is a communal component. The purpose of this section is to give an account of this aspect.

3.3.1 *Two Senses of Representation*

Our account of epistemic representation so far is very liberal. To borrow Contessa's example, someone could use Rutherford's model of the atom to represent a hockey puck sliding on ice. Or someone could use a map of Mexico City to navigate in New York City. These representations would not be accurate, but both would count as epistemic representations of the hockey puck and New York City respectively, in this first sense of representation, so long as the user assumes or pretends that they are accurate.

This way of putting things looks fairly counterintuitive, if not contradictory: how can the same vehicle be a representation of both Mexico City and of New York City? However, note that in these examples, the mention of Rutherford's model of the atom and a map of Mexico City was not explicated. In practice, the vehicle used were, presumably, inscriptions on paper. Why think that these inscriptions correspond to Rutherford's model of the atom, or to a map of Mexico City respectively? It would have been as appropriate to simply describe these examples as uses of inscriptions on paper, and perhaps to observe that similar inscriptions were once used to represent atoms or Mexico City, and no problem would have ensued. Yet, our initial formulation seems to make sense, and so there seems to be a sense in which inscriptions on paper can be considered instances of Rutherford's model, or a map of Mexico City, even if not used as such, and even if not used at all, actually: the map of Mexico City could have stayed in the user's pocket; it would still have been a map of Mexico City.

My interpretation of this is the following: the presentation of these examples mixes two senses of representation. One sense concerns the way the vehicle is used, and the other concerns its status from the point of view of the community. The same vehicle can be used as a map of New York City while at the same time having the communal status of being a map of Mexico City without any contradiction.

We should therefore distinguish between two senses of representation. This distinction follows the one proposed by Grice (1989) between utterance (or speaker) meaning and expression meaning (see Sect. 2.4.3). According to Grice, utterance

meaning corresponds to the belief that a speaker intends to convey to an audience in a particular context, while expression meaning corresponds to the appropriate use of words in general: the latter can be analysed as norms constraining the former. Similarly, the two-stage account of representation distinguishes between contextual representational use and communal representational status. The contextual use of a map concerns the fact that a given object is used by someone to navigate a city (but no particular constraint is put on the object used and its reliability) while the representational status of a map concerns the fact that an object is rightly identified, within a community, as a representation of a particular city (but it need not be used by anyone). The former sense, associated with the attitudes of the user, is the sense of representation analysed in the previous section. The latter, associated with communal norms, is the object of the present section.

3.3.2 *Licensing*

According to the two-stage account of representation, the two senses of representation just presented are not independent, in that communal status is defined by reference to contextual use (but not the other way around). More precisely, an abstract symbolic structure is a representation of a target or type of target in this second sense if it is considered optimal, in a community, to use the vehicles instantiating this structure as representations of this target or targets of this type in the first sense. Such use is licensed by the community, in the sense of Boesch (2017) (see Sect. 2.4.2). Let us say that such use is considered *relevant* in the community.

We have seen that the first sense of representation is typically associated with representational activities in experimentation. This second sense can be associated with abstract representational activities when theorising: the development of models and theories that could, in principle, be applied in many contexts. This includes, for example, the presentation of a scientific model in a classroom. In these abstract activities, models do not in general represent concrete objects with which the user can interact directly. They seem to represent types of objects or fictitious objects. When concrete objects are used, they serve a heuristic, pedagogical or illustrative purpose. According to the two-stage account, the aim of these activities is to present or develop communal norms of representation that constrain particular uses. These norms take the form of what I shall call *general models*, which are abstract entities. General models are prescriptions concerning the appropriate way of representing concrete targets. Theoretical models are examples of general models. In effect, a physicist constructing a model of the hydrogen atom in her office is developing potential norms concerning the right way of modelling particular, concrete hydrogen atoms (or hydrogen gases) in experimental contexts, and a teacher presenting this model in a classroom is telling her students how concrete hydrogen atoms ought to be represented. These norms become effective once accepted by the epistemic community.

Norms of representation are not arbitrary conventions. They serve an aim, which is, I presume, to ensure that representational uses conforming with them will be accurate. The success or failure of particular uses conforming with these norms should affect their acceptability. If a map of Mexico City does not enable its users to go from one point of the city to another, then the map should probably not be licensed as a map of Mexico City. Perhaps other factors than accuracy of uses are involved in the assessment of these norms, and perhaps in the case of science in particular, a more ambitious aim, such as truth, is involved: the axiological debate of scientific realism will precisely bear on this question. But at the very least accuracy should be involved as a criterion for an epistemic representation to be licensed.

This deflationary interpretation of epistemic representations in general, and theoretical models in particular, as conveying norms of use, could be seen as biased in favour of an instrumentalist picture of science. I do not think that this is the case: the point made here is that theoretical models do play the role of constraining particular uses. Perhaps they can receive a more substantive interpretation, in realist terms for example, either at the general level or at the level of particular uses (I will sketch such a proposal in Chap. 6, when discussing scientific realism). Perhaps the scientists constructing these models or using them interpret them in this way, and perhaps this interpretation is justified. All these remain open questions, and they will be addressed later in this book. But it could be doubted, *a priori*, that all scientists are realists, or that the ones that are instrumentalists fail to represent. All these questions are secondary from a pragmatic perspective, since a realist interpretation of this kind is not essential to the function played by these models in the economy of science, which is what interests us here. So, it makes sense to understand a theoretical model as providing a set of norms concerning the appropriate way of representing target systems of a certain type, and *then*, in a second phase, to wonder if the success of these uses should be explained by some kind of correspondence between these models and the world, or if truth is part of the motives of science.

In the case of science, higher-level norms constraining the form that any acceptable theoretical model could take (in particular, that they should respect the general laws and principles of the theory) are likely to be involved as well, on top of the norms constraining legitimate uses of particular models. Being embedded in a higher-level unifying framework is what distinguishes theoretical models from other general models. However, not all epistemic representations take place within a theoretical framework, and yet, they can be associated with norms concerning their appropriate use, which is what I am interested in in this section. The status of theories will be addressed in Sect. 3.4.

3.3.3 *General Models and Indexicality*

Before analysing the notion of relevance involved in norms of representation, let us first focus on their vehicles: what I have called *general models*. One can think of

them as exemplars or prototypes showing how targets of a certain type ought to be represented in particular contexts.⁶

The main difficulty for understanding the status of general models is that the targets and properties of interest in an experimental context are often denoted by ostentation, from a particular position. The pendulum that an experimenter could represent in a laboratory does not have a proper name, at least not one that should figure in a physical theory. A general norm cannot refer to these contextual objects and properties. It should rather refer to a *type* of object: *the* pendulum, or *the* hydrogen atom in general.⁷ But the idea that a model could represent a type of object could seem a bit mysterious: should this type correspond to a platonic universal living in an abstract world? At least it seems to commit us to a particular ontology that an empiricist might like to avoid.

The best way to address this issue without inflating our metaphysical commitments is to borrow from philosophy of language the notion of indexicality, as analysed by Kaplan (1989). Indexical terms like “I” or “now” do not have a referent outside of a context, but according to Kaplan, they have a “character”, which is a function from context of locution to content. They thus acquire a content or referent in context. For example, the character of “I” specifies that it refers to the speaker, and the character of “now” to the time of the utterance. Sometimes salience is involved: “the cat” refers to the salient cat in context. Expressions containing indexicals, such as “I am hungry”, also have a character, and their particular utterances have content (in this case, that the speaker is hungry).

General models, the function of which is to convey norms of relevance, can be conceived of as indexical in this sense. They have a character. For example: the model represents the salient pendulum; the symbol “O” (the origin of coordinates) refers to the centre of mass of the salient pendulum, etc. In context, these components acquire concrete referents that can be the object of concrete purposes, and the inferences allowed by the model become interpretable in terms of a concrete object. Outside of a context of use, we can only imagine how these inferences would be interpreted. When we do so, the model has a fictional target. Imagining a fictional target is probably what theoreticians do when developing their models, which is why fictionalist accounts of representation can boast about being able to account for the way scientists talk about their models (Sect. 2.4.1). But the genuine representational status of general models is given by their character: the way they could be applied in various contexts.

We can then understand how a general model acts as a norm licensing particular uses. A general model can be characterised by a symbolic structure described, for example, by equations expressed in a theoretical vocabulary. Consider, in a given

⁶Note that a general model gives us permission to represent a target in a certain way, which does not preclude that other ways, conveyed by other general models, are permitted as well: the present account does not preclude pluralism.

⁷The exception to this is when the target does have a proper name: for example, the sun, or the solar system.

context, all possible interpretations of this structure in terms of the context, that is, the possible mappings between its symbols and properties of interest. From a communal perspective, only some of these interpretations make sense, and in some cases, perhaps none of them make sense. Mapping the position of the pendulum in the model to its colour is not a correct use of the model. Interpreting Rutherford's model of the atom in terms of a hockey puck cannot be considered a legitimate use of this model either within the scientific community, whatever the interpretation, because a hockey puck is not an atom.

Let us say that the character of a general model specifies, for any context, the possible denotations of its symbols, that is, the (possibly empty) set of interpretations of its structure in terms of this context that are licensed by the community, or *relevant*. Once interpreted, this structure becomes what I have called in the previous section an abstract interpreted model, so in effect, the character, or relevance function of a general model, gives, for any context, the set of abstract interpreted models instantiating this structure or part of this structure that are licensed. Let us call a general model the combination of a symbolic structure and a character.

Definition 3.4 General model := $\{S, r\}$ where S is a modal symbolic structure and r a function taking a context C as a parameter and returning a set of abstract interpreted models $\{m\}$ each instantiating a sub-structure of S interpreted in terms of C .

In sum, a general model is characterised by its structure and by its set of licensed representational uses. The relevance function ensures that a model of the hydrogen atom does represent an hydrogen atom, and not a hockey puck, and that the position of the electron in the model represents the position of an electron, and not any other property. This notion can easily be extended to concrete vehicles, so as to account for the fact that, say, a sheet of paper with particular inscriptions *is* a map of Mexico City. In this case, the concrete vehicle is also associated with a relevance function.

Note that it is enough for the interpreted model to instantiate a sub-structure of the general model, because in particular contexts, users might not be interested in the complete structure. For example, a user could use the general model of a pendulum, but assume, because of empirical inputs that are specific to the context, that the initial position of the pendulum has a particular value. Dynamical parameters will also be fixed by the context. In this case, the modal structure of the general model comprises all possible initial positions and dynamical parameter values, but the structure of the interpreted model only contains one possible initial value. In this sense, we can say that the interpreted model makes the general model more precise, in a way that is similar to Giere (1999)'s account of the relationship between abstract and concrete models (see Sect. 2.3.1). Contexts are also bounded in space and time and denote discrete properties in finite number, while a general model can be infinitely extended and continuous.

These observations agree with our remarks from the previous section on the fact that an interpretation is akin to a projection on a perspective. We could say that a general model synthesises various possible perspectives on a type of object associated with potential contexts, and sometimes an infinity of them.

I assume, for the sake of generality, that more than one application of a general model could be licensed in a given context, although I presume that in general, at most one will be licensed. It could also be that none are licensed if the general model does not apply to the relevant target in this context, in which case the function r returns an empty set. For example, Rutherford's model of the atom will have no licensed interpretation if the context refers to a hockey puck. Thus, the character of a model delimits its domain of application, which is the set of contexts for which the relevance function returns a non-empty set.

This domain of application can be associated, intuitively speaking, with the type of object that the model represents. For example, the model of the pendulum represents a certain type of mechanical system, which can be characterised by the features shared by all contexts in which the model has licensed interpretations: the general characteristics of pendulums. However, it is important to remember that these contexts are associated with particular purposes, so that the type associated with a general model does not necessarily correspond to a factual category of objects or to a natural kind. It is more accurate to associate it with a category of *purposes*, or type of activity, which, incidentally, can concern a category of objects (see Knuuttila and Boon 2011 for a similar view). Pendulums are classified as such from a perspective which focuses on mechanical characteristics for instance, and this perspective implies a certain type of activity. Similarly, a map can be associated with a type of activity, such as navigating in a particular city.

It will be useful, for notational purposes, to introduce a new function that specifies the set of applications of a model that are licensed. This function, let us call it *relevance*, takes a general model and a context as parameters, and returns a set of abstract interpreted models.

Definition 3.5 $relevance(M, C)$: specifies the set of abstract interpreted models $\{m\}$ that correspond to licensed uses of a general model M in a context C .

This function is introduced for the purpose of simplifying notations, but it brings nothing new: it amounts to applying the function r of the general model M to the context C , where M and C are its parameters.

In the following chapter, empirical adequacy will be defined in terms of the *relevance* and *accuracy* functions.

3.3.4 Relevance

Let us say a bit more about the notion of relevance.

According to the two-stage account of representation, general models are indexical, and their main function is to convey norms of representation, that is, to specify a set of relevant uses for a given symbolic structure. Whether an interpretation is licensed or not can depend on many things.

For example, whether or not a particular use of a city map is licensed will typically depend on a correct interpretation of its symbols: such symbol should be

interpreted as a street, as a metro station, and so on. In the case of science, this will correspond to a correct interpretation of theoretical vocabulary: the position of an electron in a model can be mapped to the result of a measuring apparatus that was designed to measure particle positions, but not to an apparatus that measures the strength of the electric field or whatever other property. Since we are assuming that the contexts in which theoretical models are interpreted are theory-laden entities, interpreting the theoretical vocabulary correctly simply means that a position in the model should be mapped to a position in the context, or to say it differently, that a particular symbol of the general model is a position, and not any other property.⁸ A map should also be oriented correctly for appropriate uses, and I assume that this aspect boils down to a correct interpretation of symbols (the North symbol, for instance).

Other norms of relevance concern factual knowledge, postulates or hypotheses concerning particular types of targets, including what is often referred to in the philosophy of science literature as auxiliary hypotheses. However, these norms are conveyed differently. A map will be licensed by the community for being used in Mexico City because it contains the relevant features of the city. Similarly, in the case of science, a model of the solar system should contain eight planets, or a fluid should be modelled using Navier-Stoke's equations. This means that a model with only six planets will have no legitimate interpretation in a context where the solar system is represented. Norms of this kind do not concern the mapping between symbols and contexts, but rather the type of system that is being represented (or the particular system if it has a proper name), or the type of activity for which the model can be used. If the target of representation is not the right one, or not of the right type, then no interpretations will be licensed. If it is the right one, then the model conveys, by means of its structure, appropriate ways of representing this target or type of target. As said earlier, this does not prevent other models to be licensed as well, but we can assume that the fact that a model is licensed precludes incompatible models to be licensed as well, so that the model implicitly conveys norms concerning the right way of representing targets of a particular type for particular purposes. A relevant type of target is specified by the context by means of a fixed property (see Sect. 3.2.1), which can be simple ("being a fluid") or correspond to a more complex description, in which case the role of the general model is also to define a certain category of objects or activities.

The mention of activities is important, in particular in science, because whether or not an interpreted model is relevant also depends on contextual aspects that do not have to do with the represented object, but with the purposes of the user, in particular, relevant properties and expected degrees of precision. A metro map can be used for some purpose, such as going from one station to another, but not for another, such as evaluating the distance between stations, because it distorts distances. In science, some idealisations, such as ignoring air friction in a

⁸Questions related to the operationalisation of theoretical terms will be addressed in Sect. 3.4.2. As I will argue, norms of operationalisation are not conveyed by theoretical models.

mechanical model or considering that a gas is infinitely extended, could be relevant in a context, but unacceptable in another, even though the target system is exactly the same, only because levels of requirement are higher concerning some accessible properties. Whether or not neglecting friction is acceptable given expected degrees of precision depends on theoretical considerations, so it cannot be captured by a correct interpretation of symbols alone. But this sensitivity to purposes can be readily captured by the fact that a character is a function that takes the context as a parameter.

In sum, the relevance function mixes three potential aspects:

1. the correct interpretation of the symbols of the general model,
2. factual knowledge or theoretical postulates associated with a target or type of target, and
3. acceptable idealisations in a given context.

Relevance does not imply accuracy. The fact that the epistemic community considers that a particular use of a model is correct does not mean that this model is accurate. At this stage, misrepresentation is also possible. However, as noted earlier, relevance is likely to be sensitive to general empirical inputs, in particular if norms associated with factual knowledge or postulates concerning a type of target and activity are involved. A map of Mexico City is licensed as such by a community in so far as it has proved its reliability. The map can contain mistakes. However, it would be surprising if it got nothing right about its target: why was it licensed in the first place? The right way of modelling phenomena of a certain type presumably depends on a certain accuracy of past uses of the norm, that is, on a certain experience accumulated within the community. In this sense, accuracy is not a condition for a model to represent its target, but one could say that a certain degree of accuracy is a by-product of the licensing process, so that a model representing a target in the communal sense will often be accurate.^{9, 10}

However, in the case of science, the domain of application of a general model is sometimes extended to yet unexplored contexts, and there is no guarantee that the model will still be accurate in these new contexts. So, the fact that norms of relevance are responsive to empirical inputs and that they build on past experiences does not preclude cases of misrepresentation.

⁹This could shed light on the controversy between similarity-based and user-centred accounts of representation (see Sect. 2.3.2).

¹⁰Note that according to this account, idealised models only “misrepresent” their targets in the sense that the range of contexts for which the model is licensed is deliberately limited. The idealised model does not misrepresent when it is properly used, because it affords inferences that lead to true conclusions. It *would* misrepresent if it were used in more fine-grained or more comprehensive contexts to represent the same target. Although this contradicts a common characterisation of idealisations in terms of deliberate misrepresentation, I think that this consequence of my account is not problematic. A map of metro stations that distorts distances does not really misrepresent its target once it is acknowledged that it does not aim at representing the distances between metro stations in the first place.

Incidentally, norms of relevance are in principle revisable. The best way of applying a general model, of delimiting its scope and of classifying phenomena that must be modelled differently, can be learned by experience. As we will see, part of the axiological debate on scientific realism concerns the values that drive the revision of norms of relevance.

3.3.5 *Modularity and Composition*

Our account of representation is incomplete as it stands, because of a feature of general models that is particularly salient in science: their ability to combine into more complex models. Simple models can be used to construct more complex models, and in these cases, the simple models are not directly interpreted in terms of contexts, or not always. For example, a biologist interested in the progression of a virus in human organisms will have at her disposal various licensed models representing the different steps of this progression: the virus binding to proteins and entering the host cell, the transcription of the viral genome, etc. (I owe this example to Mitchell and Gronenborn 2015). The simple models used in this complex representation can be separately interpreted in terms of particular experimental contexts, and validated in these contexts, for example, when the characteristics of a single protein are observed by chromatography. However, the composite model will typically be interpreted in contexts where the simple models are not directly interpreted, for example when monitoring the effect of a medicine on the progression of a virus. These contexts do not overlap. This is an example from biology, but it should not be difficult to find similar cases in physics or other disciplines.

What is important is that in such cases, the relevance functions of the simple models are not involved directly (or at most, only partially), since the features of these simple models are not the focus of the users' interests. To be sure, the fact that a simple model purports to represent a certain kind of protein and that its symbols are interpreted in a certain way matters for the construction of a composite model. However, these norms of relevance are, so to speak, absorbed in the norms of relevance of the composite model.

This does not affect the previous remarks. It only implies that a new kind of norm plays a role in representation, which I shall call *compositional norm*. A compositional norm can be conveyed by an abstract model describing a "pattern" (the progression of viruses in general), the details of which are filled in by simple models to yield a composite model. I suppose that there is some flexibility in the construction of composite models. In virtue of this flexibility, compositional norms do not preclude the existence of norms of relevance that are specific to the composite model. For example, it is likely that among the many legitimate ways of combining simple models to represent the progression of one particular virus in an organism, only some are eventually licensed by the epistemic community. In this case, the composite model incorporates norms of relevance of its own, conveyed

by the chosen combination (the simple models that were used and the precise way in which they were combined). These norms are intuitively associated with factual knowledge about the target of representation: that such protein is involved in such a way in such phenomenon, and these norms do not supervene on the relevance and compositional norms that were used in the construction of the composite model.

Interestingly, the fact that the simple models are independently licensed in their own contexts of application seems to confirm the legitimacy of the composite model. Conversely, the fact that various combinations of a simple model are licensed could confirm the legitimacy of the simple model. This aspect will be addressed in the next chapter. However, just like for other norms, a realist interpretation of composite models in terms of material composition is not required, in so far as we are only concerned with the pragmatics of representation.

3.3.6 *Explanations*

This account of representation is deflationary in spirit. The main idea is that epistemic models convey norms of representation, which can then be applied in context in order to make inferences about concrete targets of representation. This deflationary aspect might seem unsatisfactory to some readers. In particular, one could object that general models do more than just licensing inferences in particular contexts. The scientific models presented in classrooms, for instance, give us a certain understanding of the world, notably because they *explain* phenomena of interest.

We cannot reasonably address the debate on scientific realism without saying a word about explanations, because of the inextricable link between the two (this link will be explored in Chap. 6). It is generally accepted that one of the main roles of scientific theories and models is to explain natural phenomena. The point raised here is that one could wonder how the notion of explanation fits in the account of representation presented in this chapter.

Explanations are often thought to be part of a family of interrelated concepts including laws, causation and counterfactuals that are modal in character. Thus, among the major theories of explanation are accounts in terms of covering laws (Hempel and Oppenheim 1948), unification (Kitcher and Salmon 1989), causal mechanisms (Salmon 1984) or counterfactuals (Woodward 2003; Saatsi and Pexton 2013). Evaluating these various accounts is beyond the scope of this work. However, it is reasonable to assume that whatever theory of explanation one adopts, general models, as defined in this section, actually have all the necessary features to be explanatory.

Indeed, the relevance function of a general model brings unity to its various possible applications (this notion of unity will be analysed in Sect. 4.3.3). Although this might not imply a fundamental law of nature, it makes sense to assume that proposing a general model is tantamount to proposing an experimental, or phenomenological law (this will be defended in Sect. 5.2.2). As for causation

and counterfactuals, I have noted that interpreted models generally have a modal structure, and this makes them fit for representing causal processes and allowing for counterfactual reasoning (I return to this aspect in Sect. 4.2.6). The possibility to combine simple models into more complex ones could also be relevant to the notion of mechanism, on which there is a growing literature. The present account of epistemic representation does not purport to be an account of the psychological aspects associated with explanations and understanding, because I think that these are derivative rather than essential features of representation, but there is no principled incompatibility between this account and the idea that models are explanatory.

The theories of explanation mentioned above do not exhaust the landscape of philosophical positions on the topic. Some authors have argued that all these theories are unable to account for the sensitivity of explanations to purposes and contexts. For example, we might explain a death following a car accident by invoking either multiple haemorrhage, negligence on the part of the driver, a defect in the brakeblock construction or the presence of tall shrubbery at that turning, depending on our focus (Hanson, cited by van Fraassen 1980, p. 125). Or we might explain the length of the shadow of a tower by invoking its height and the position of the sun in one context, and explain the height of the very same tower from the length of its shadow if the tower was designed to cast a shadow of a certain length in another context (van Fraassen 1980, ch. 5). According to van Fraassen, an explanation is an answer to a “why question”, and the answer depends on the contrast class or (as in the examples above) on the relevance relation implied by the context. This relation of relevance notably affects what counts as a fixed background for counterfactual reasoning. Van Fraassen claims that “scientific explanation is not (pure) science but an application of science. It is a use of science to satisfy certain of our desires.” This relation to a context is all there is over and above the descriptive content of the explanation given by the theory, and this allows van Fraassen to claim that explanatory power simply derives from the adequacy of the theory. Given the contextuality of the present account of scientific representation, where the context encapsulates the interests of agents for particular applications of a model, including fixed and variable properties (Sect. 3.2.1), it should be clear that this account also has the resources to address this aspect of explanations, and in a way that is more closely associated with scientific activities than what van Fraassen suggests. Users of a scientific model assume that the model would apply to different contexts, so the unificatory aspect is still present.

Whether or not an explanation needs to be true to be a good explanation is at the centre of the controversy between empiricists and scientific realists. I will defend in this work that the interpreted models of our best theories can be considered accurate including for merely possible situations, where the possibilities considered are mind-independent. I will also defend, in an empiricist spirit, that an explanation only needs to be empirically adequate in this modal sense to be acceptable. The modal component of empirical adequacy that a modal empiricist puts forth makes this assertion more plausible, given the modal characteristics of the family of concepts just mentioned. For a modal empiricist, the modality of counterfactual inferences

can be interpreted “literally”, in terms of real possibilities, and the same goes for causal assertions and experimental laws. This means that for a modal empiricist, the link between the adequacy of a model or theory and its explanatory power is quite straightforward.

One could have the impression that this account is still insufficient, perhaps even as an account of representation in general, because scientific models are often interpreted in a more pictorial, or perhaps metaphysical way. Consider, for example, the model of the ideal gas in statistical mechanics, which depicts gases as being composed of molecules bouncing off one another like billiard balls. It could seem like this model is more than a set of potential mappings between aggregate properties of its mathematical structure and accessible properties of concrete targets (typically, mean velocities and temperatures). It depicts gases in a certain way that is not exhausted by these potential mappings. One could also have the impression that this depiction participates in the explanatory power of the model.

Perhaps such pictorial aspects are often involved in representation, but in so far as they do not affect our interaction with concrete target systems, I would say that they merely play a heuristic or psychological role, and that they do not concern us here. The present account only aims at capturing the pragmatics of representation in a minimal way. In any case, as said earlier, one is free to complement this account with a realist notion of interpretation. But such a notion is not required for representation to take place, and I would say that it is not required for explanations either. For example, I would argue that an explanation of the emission spectrum of a material from a quantum mechanical model of the atom is the same explanation whether we interpret the wave function as a nomological entity relating to particles or events, or as an ontological entity such as a multi-field. In sum, one’s preferred metaphysics does not affect the status of scientific explanations. This point will be discussed in Chap. 6. In the meantime, let us simply assume that the present account correctly captures the pragmatic aspects of representation, including when it comes to explanations.

3.4 The Norms of Representation in Science

We have now detailed the two stages of our account of representation: the first stage is formalised in terms of interpretation, which is a mapping between a symbolic structure and contextual properties of interest, and the second stage is understood in terms of communal norms constraining these interpretations, where these norms are conveyed by general models associated with functions from context to licensed interpretation. We have seen that simple general models can be combined into more complex ones, and this process is also constrained by communal norms conveyed by more abstract models.

So far, the account of representation described could be applied to any kind of epistemic representation. The aim of this section is to focus on scientific theories in particular, and examine the norms that play a role in the application of these theories.

I do not think that there is a specific problem of scientific representation, in the sense that what constitutes the representation relation in science would differ from what constitutes epistemic representation relations in general. In this respect, I agree with Callender and Cohen (2006) that the problem of scientific representation is merely an instance of the demarcation problem (see Sect. 2.4.1), that is, the problem of understanding which kinds of epistemic vehicles are considered good or bad for the purposes of science. Representational vehicles across different scientific disciplines, such as economics, psychology, biology or physics, and different activities, such as prediction, technological development or foundational theorising, are various. Many types of vehicles used in science, such as maps or diagrams, are also used outside of science. The odds of finding a common characteristic that is unique to science seem low. The idea that science is an institution characterised by shared values, in light of which some general models can be considered good and others bad, seems much more promising to me.

Having said that, it seems reasonable to believe that one of the characteristic values of science is the search for unifying frameworks, or theories, and that it is possible to say a bit more about this aspect. This means focusing on one particular type of scientific vehicles: theoretical models. The point is not to characterise how these particular models represent, because there is nothing specific to them in this respect, but to characterise the theoretical frameworks in which they are embedded, and the norms of relevance associated with them. I will focus here on two related aspects: theoretical laws and principles on the one hand, and the interpretation of theoretical vocabulary on the other. Finally, I will say a word about the notion of layers of representation.

3.4.1 *Theoretical Unity*

I take theoretical models to be general models, with the further characteristic of being incorporated in a more general framework, the theory. Following the semantic conception of theories, we can simply say that theories are families of general models.

Each theoretical model comes with a relevance function that constrains its possible applications. As noted earlier, these relevance functions would typically encapsulate factual knowledge or postulates that are specific to a target or type of target (or rather, the fact that some structures are relevant for some types of targets conveys this factual knowledge). This ensures a certain autonomy between models and theories, as observed by Morgan and Morrison (1999) (see Sect. 2.3.1). However, as noted in Sect. 2.2.2, theoretical models share a common vocabulary, and they are organised by laws and principles. A theory is not a disparate collection of models that have nothing in common. For this reason, one might want to characterise theories by providing further norms that specify the set of models that are acceptable as models of the theory, as well as their general interpretations in terms of potential contexts. These norms would typically correspond to the general

laws and principles of the theory, and their relations with the models of the theory could be specified in Tarskian terms: only the models whose structure satisfies the laws of the theory are models of the theory. In addition, norms governing the combination of simple models into complex ones could be considered. In some cases, these compositional norms might already be implied by the laws and principles of the theory that any model must respect.

Since the interpretation of the symbols of a general model is conveyed by its relevance function, and since general laws are expressed in terms of these symbols, one could view theoretical norms as *meta-norms* constraining acceptable relevance functions for particular symbolic structures. For example: a symbol can be interpreted as a position only if it respects theoretical constraints with regard to positions in general (their relation to velocities, masses, forces, etc.).

I have claimed earlier that general models are explanatory. One can think of meta-norms as giving the form that explanations must take in a particular field of inquiry built around a particular theory. For example, in Newtonian mechanics, explanations should involve forces. In the theory of evolution, explanations should involve selection mechanisms. Perhaps, drawing inspiration from Giere (1999), norms and meta-norms could be combined into a hierarchical picture, with various levels of abstraction for norms, all these norms being conveyed by more or less abstract general models (see Sect. 2.3.1). The “character” of more abstract models could be a function from *types* of contexts to more concrete models. Some subdisciplines, for example, evolutionary psychology taken to be a subdiscipline of biology, are based on general hypotheses (the controversial massive modularity of the mind in evolutionary psychology) which seem to boil down to constraints on the form that explanations must take in this particular field of inquiry, on top of the constraints of the fundamental theory. One could think of Navier-Stokes equations as giving the form that explanations must take in fluid mechanics, and the same rationale could be given for more specific theoretical models, which would constrain the form of contextual explanations by means of their relevance function.

A problem with this characterisation of theories in terms of meta-norms of relevance is that as noted in Sect. 2.3.1, some idealised models distort theoretical laws or are based on more than one theory. A model of the solar system where the sun is fixed in its reference frame does not strictly satisfy the laws of Newtonian mechanics, because planets should also exert a gravitational influence on the sun, but this simplification is often used. However, this is not so much a problem, at least for an empiricist, assuming that these idealised models generally make the same coarse-grained predictions as a model that does respect the laws of the theory. This means that the accuracy of an idealised model is apt to confirm the accuracy of the corresponding model of the theory. Meta-norm infringements do not matter if there are good reasons to think that they do not make any practical difference.

One could be tempted to downplay the importance of higher-level theoretical laws, and think of theories in terms of disparate models only. However, there are good reasons to assume that such meta-norms exist. One reason is that it ensures that the theory can be extended to new types of phenomena. Sometimes, scientists apply a theory to hitherto unexplored types of phenomena by constructing a new

model. As we will see later in this book, the success of such extension plays a central role in the debate on scientific realism. If there were no rule constraining the construction of these new models, then successful extensions of the theory would not be indicative of any kind of success at the theory level. However, the history of science shows that these extensions play an important role in the general acceptance of theoretical frameworks (see for example Zahar 1973).

A second reason to accept these meta-norms is that they can provide identity conditions for theories. I have mentioned earlier that norms of relevance are revisable when they concern factual knowledge or domain-specific postulates. Presumably, new norms of relevance become licensed when new models are constructed for new applications of the theory. This means that a theory, characterised as a family of general models, can evolve, but it remains the same theory. Having fixed meta-norms at the theory level constraining the set of legitimate general models is important for explaining in what sense it is still “the same theory”, and in what sense updating the way particular target systems are represented in Newtonian mechanics is not the same thing as switching to a completely different theory such as general relativity. This distinction seems important to account for the way science functions.¹¹

At this point, we could go as far as claiming that a theory is just a set of meta-norms, and that factual knowledge or posits concerning particular domains of application, that is, the norms conveyed by theoretical models, are not part of the theory itself. This way of thinking is quite common in philosophy. For example, domain-specific posits are often called “auxiliary hypotheses”, implying that they are not really part of the theory itself. This would mean that the collection of models that characterise a theory (any model satisfying these meta-norms) would be much larger than the collection of models actually licensed by scientists. In this context, having an intermediate level of general models between theories and interpreted models could even seem superfluous: one could just as well say that a theory is a very abstract general model licensing many interpreted models in particular contexts. However, I believe that it makes more sense to construe a theory as a family of licensed theoretical models with specific members that incorporate domain-specific norms, and to take meta-norms to be mere identity conditions for

¹¹Having said that, I should note that I do not want to be too sanguine about the idea that theories have strict rather than vague identity conditions. French (2017) has argued against this idea. Identifying different theories or research programmes such as Newtonian mechanics, general relativity or the theory of evolution is, at least, a convenient way of talking about the product of science, perhaps in the same way as identifying musical styles is a convenient way of talking about the product of the music industry. Theories are probably more easily delineated than musical styles, by means of meta-norms that persist even when local models evolve. However, these meta-norms might also be flexible. The picture is also complicated by the fact that some models are constructed using more than one theory (style fusion does not only happen in music!). I will not give detailed answers to questions such as “would Newtonian mechanics still be the same theory if the law of gravitation were different, or if it were expressed in geometrical terms instead of forces?” in this work (the question of theory equivalence is briefly addressed in Chap. 8). Although I will talk as if theories and their models could be strictly identified, the notion of meta-norm is more important for the purpose of this book than the notion of theory.

the theory.¹² In this context, we could distinguish between licensed models and candidate models. The former are licensed models of the theory, and the latter are not, but they still satisfy the meta-norms of the theory, and they could be used for application of the theory to new domains.¹³

An argument for this approach will have to wait for the next chapter. Until then, let us just say that it makes it easier to characterise the aim of science, taking theories to be the main product of scientific activity, because domain-specific theoretical models are certainly an important product of scientific activity. Understanding them as merely instrumental in confirming the meta-norms that characterise the theory seems too reductive.

In any case, we can remain neutral in this matter for the moment by simply sticking to our understanding of theories as collections of general models, without presupposing anything concerning the way these models are organised.

Definition 3.6 Theory:= collection of general models $\{M\}$.

The actual domain of application of the theory is fully specified by these models and their relevance function. Including candidate models as well gives us a larger domain, which we could call the potential domain of application of the theory. One could characterise the cognitive content of the theory by the *relevance* function defined above, restricted to the models of the theory, since it tells us exactly how a theory ought to be applied in any context.

3.4.2 The Interpretation of Theoretical Terms

One specificity of the characterisation of the range of acceptable models of a theory by means of general laws and principles is that these laws and principles are expressed in a theoretical vocabulary that is shared by all theoretical models. For example, the laws of Newtonian mechanics relate forces, masses and the second derivative of positions. Without this common vocabulary, not only could models not be organised by general principles, but they would be pure structures, without any connection to experience. So far, I have not dealt with the way this vocabulary is interpreted in particular contexts, that is, the way it is transcribed into observations and manipulations of target systems. I have simply assumed that contexts are theory-laden entities, and that theoretical structures are mapped to the context. It is now

¹²This is analogous to Lakatos (1978)'s distinction between the core of a research programme (a sequence of theories sharing this core) and its protective belt.

¹³Finer distinctions can be made: for example, a complex model combining licensed models looks like a better candidate than one that combines non-licensed ones. Some candidate models are incompatible with licensed models. In a hierarchical account, the place of candidate models in the hierarchy could also matter.

time, before concluding this chapter, to say a bit more about the relation between theoretical terms and experience.

As remarked in Sect. 2.3.3, the transcription between theoretical terms and observations or manipulations can be fairly complex. It can involve statistical techniques used to extract relevant information from brute data, or data clean-up considering contextual factors. It can involve auxiliary theories from which the outcomes of measuring apparatus are interpreted, as well as practical knowledge concerning the appropriate use of these apparatus, and a conceptual ability to recognise relevant patterns. Most factors involved in scientific experimentation go far beyond the local context (for example, knowing that such apparatus was manufactured by a trustworthy company, that it was bought recently and should work properly), and attempting to codify them all in terms of systematic rules of application for the theory is hopeless.

Nevertheless, it seems reasonable to assume that all these aspects follow communal norms constraining particular uses, such as legitimate operationalisations for quantities, in the same way that the choice of a theoretical model to represent a particular target follows communal norms. The relevant community in this case is, I suppose, the community of experimenters. These norms are directed towards particular values: the stability and reproducibility of experimental results, and the unification of various operationalisations when they give similar results. As before, we do not need to presuppose that any version of realism (such as entity realism) is implied, as long as these norms capture the surface features of experimental practice in science. Whether experimental successes warrant the belief that theoretical terms refer to natural properties is a secondary question from a pragmatic perspective. This question will be addressed later in this book.

These experimental norms can perhaps overlap with theoretical norms in case auxiliary theories are used in experimentation, for example, when they are involved in understanding the functioning of apparatus. At least both types of norms should be coherent. However, the way they affect each other is not necessarily unidirectional (see Sect. 2.3.3).

I mentioned earlier that theoretical models can be combined into more complex ones, and it could be thought that this notion of combination is enough to account for experimental contexts from a theoretical perspective, without assuming separate experimental norms. The idea would be that the models that are ultimately confronted with experience are complex combinations of the kind described earlier, typically a combination of a model of the target system and a model of the instruments involved in measuring this target. This idea is idealistic, and I believe that we should keep experimental and theoretical norms distinct for at least three reasons.

Firstly, scientific instruments need not be entirely modelled in a theory to be considered reliable (see Sect. 2.3.3). Experimental techniques often evolve independently of theories, and they can survive theory change, so these two types of norms have a certain autonomy.

Secondly, if theories only were involved in the mapping between theoretical terms and experience, we could fear an experimental regress: the theories would be

confirmed or discarded by their own lights. A quantity would count as a temperature in a data model only if it respects the laws of thermodynamics, and thermodynamics could never be rejected by any experiment, because all data models would respect its laws. But in general, the circularity is not complete (Franklin et al. 1989). Even when measuring instruments are modelled in a theory, these theoretical models must be connected to our observations, and this is the role of experimental norms.

The third and final reason for maintaining a distinction between theoretical and experimental norms is that the interests of model users rarely lie in the position of needles or in the value of numbers displayed on a screen. They are interested in theoretical quantities, and theoretical models can be operationalised in various ways. A theoretician developing a suitable model for a given type of phenomenon will follow theoretical considerations, perhaps taking advice from empirical inputs expressed in theoretical terms. She need not be bothered by the precise way these empirical inputs were produced, or by the various ways in which her model could be operationalised. This is a problem for experimenters. The same goes for someone applying this theoretical model for concrete purposes: particular operationalisations are contingent with respect to her purposes, which are generally best expressed in theoretical terms. All that is required is that these operationalisations are efficient.

I would say that the respective domains of application of experimental and theoretical norms are mainly determined, in context, by users' particular interests. It might sometimes be useful to model instruments in the theory for some purposes (when manufacturing these instruments, or when trying to understand their functioning), in which case theoretical norms of representation will be applied, but for other purposes, it will be enough to assume that these instruments are reliable and to follow experimental norms. The two kinds of contexts involved are distinct. The fact that users' interests can vary, and that experimental or theoretical norms can be applied at different places for a single target of representation depending on these interests (modelling or not an instrument in the theory), imposes a certain coherence between these two types of norms. This could explain why experimental norms are impacted by theories, or conversely. However, the separation between two domains of application for communal norms, one "upstream" and one "downstream" from where a particular user's interests lie, one concerning the target of representation and one its measuring environment, is always present whatever the context, so these two types of norms are distinct.

This division of labour between theory and experience makes sense because a theoretical model purports to represent a type of *phenomenon*, and not brute contextual *data* (see Sect. 2.3.3 for the distinction between data and phenomena). One could say that the role of experimental norms is to connect data and phenomena, while theoretical norms only concern the representation of phenomena (including sometimes the ones that occur in measuring instruments).

All this to say that experimental norms are not conveyed, but assumed by theoretical models, which legitimates the stance of considering contexts to be theory-laden entities that has been adopted so far. The norms associated with licensed operationalisations of theoretical terms and the norms associated with

the relevance of theoretical models for representing particular types of phenomena should be kept distinct.

3.4.3 *Layers of Representation*

Shall we say that experimental norms are conveyed by general models, just like theoretical norms, and shall we apply the same account of representation to them? I believe that we can, at least to some extent. A general model unifies various contextual uses, and similarly, theoretical terms unify various possible operationalisations, so there seem to be a relevant analogy. One type of model is involved in particular: what Suppes (1969) calls a *model of the experiment*. These models act as mediators between a theoretical model and experience. They are involved in the interpretation of brute data and in the construction of data models that are apt to represent phenomena. For example, a model of the experiment could be used to synthesise astronomical data gathered from three different telescopes located in different places, so as to describe a phenomenon in a way that is relatively independent of the precise location of these telescopes¹⁴ (Bailer-Jones 2013, pp. 170–171).

Arguably, a data model and a model of the experiment are epistemic representations of a particular target, which is an experimental configuration, and since our two-stage account of representation is not restricted to theoretical models, it should apply to them as well.

If models of experiments and data models are epistemic representations, then according to our account of epistemic representation, they should be interpreted in terms of a context. If their role is to connect theoretical terms and direct observations, then the contexts in which they are interpreted are presumably superficial ones associated with directly observable objects described in natural language instead of theoretical properties: apparatus, needles, computer screens, etc. These contexts are also presumably quite broad (as noted earlier, the manufacturer of an apparatus can be relevant). So, an interpreted model of the experiment would provide a mapping between a symbolic structure expressed in a theoretical vocabulary, which typically specifies a set of conceivable outcomes for an experiment, and a broad experimental context interpreted in natural language. As in the case of theoretical representation, only some of the conceivable states specified by the broad context are “permitted” by the interpreted model: the others would not correspond to the intended experiment.¹⁵ Some of these interpretations of theoretical terms in terms of

¹⁴Note that this is not the same as modelling the telescopes in the theory.

¹⁵We can see that in experimentation, the direction of fit is typically from the experimental context to the model: if the model is inaccurate, then the experimental context must be adjusted, not the model. This could provide another way of distinguishing the respective domains of application of experimental and theoretical norms (but there might be complications).

broad situations are licensed by the community: they correspond to the appropriate ways of operationalising these theoretical terms. They might be conveyed by *general models of experiments* that would represent types of experimental activities. In this sense, norms of experimentation could be conveyed by general models, at least in part, in full compatibility with the account of epistemic representation presented in this chapter.

I have claimed earlier that the notion of context is close, in the case of science, to the notion of model of the experiment proposed by Suppes. A model of the experiment determines the “form” of the data. This can be expressed, as proposed by Suppes, by providing a set of conceivable data models, in the same way our notion of context specifies a set of conceivable theory-laden states for a target system. What could seem puzzling is the idea that these models of experiments would play the role of a context for theoretical models, while being themselves interpreted in terms of a context. This implies several representational layers: a theoretical structure is interpreted in terms of a model of the experiment, which is itself interpreted in terms of superficial properties of an experimental configuration. Having several layers of representation is not necessarily problematic (a similar idea has been suggested by Latour (1993), and, in a more formal context, by Bueno (1997)), but one could wonder whether this leads to an infinite regress, or if at some point, a context need not be interpreted anymore: is the broad, superficial context in which a model of the experiment is interpreted an epistemic representation as well? In terms of an even broader and more superficial context perhaps? But where shall we stop?

This touches on a deep philosophical question that has to do with the relationship between representation and the world in general. Van Fraassen (2008, pp. 253–261) discusses a similar issue, and claims that the fact that a data model is a representation of its target is a “pragmatic tautology”, in the sense that claiming “It isn’t so, but I believe that it is” is not a logical contradiction, but is pragmatically awkward. This strategy can be seen as a way of stopping the regress at the level of data models. With this kind of pragmatic tautology, there is nothing more to be said in terms of relations between various *representations*, since the link is between our representations and the world directly. However, assuming our two senses of representation (Sect. 3.3.1), this claim is ambiguous. On the one hand, considering contextual use, the fact that a data model is a representation of its target does indeed look like a pragmatic tautology, in so far as contextual representational use rests on the intentional attitudes of users. But this is true of any kind of contextual representation, including when a map or a theoretical model is used, and there is nothing specific to data models here. On the other hand, considering communal representational status, the assumption that the data model is a representation of the target intended by the user can be denied if the experiment was poorly performed, so it is not a pragmatic tautology (see also Nguyen 2016).

My guess is that the regress stops a bit further, when representation is markedly mental, without any external vehicle being involved. There is a form of immediacy associated with perceptual representation that is not present when representation is externalised by means of non-mental vehicles. This suggests that in the case of perceptual representation, we have first-order representation: the mental vehicle

is no longer interpreted in terms of another representation. This is our ultimate context, and since it is private, there is no question of communal licensing. From the perspective of the community, a data model can fail to represent anything if experimental norms were not respected. But the fact that the mental representation of someone represents that towards which it is intentionally directed, accurately or not, can hardly be denied at the community level. In this sense, we could say that mental representation is licensed by default, and that the fact that it represents what it represents for its “user” is indeed a pragmatic tautology: we reached the point where nothing more can be said in terms of relations between various representations, because the link is to the world directly.¹⁶ This would mean that ultimately, as suggested by Callender and Cohen (2006), these matters would fall under the scope of philosophy of mind.

3.5 Epistemic Values and the Axiological Debate

According to the two-staged account of epistemic representation presented in this chapter, representational use involves an interpretation, which is a mapping between a vehicle instantiating or describing a symbolic structure and a context. The context can be formalised as a set of conceivable coarse-grained states or histories for a target system, which represent the interests of a user. The interpretation restricts these conceivable states to the ones “permitted” by the model. This constrains the inferences made on the target. Such use can be licensed or not by communal norms of representation. These norms are conveyed by general models, which are symbolic structures associated with a function from context to licensed interpreted models. They constrain the way particular types of targets ought to be represented. Simple general models can be combined into more complex ones, and they can be explanatory. Scientific theories are families of models. They are identified by meta-norms constraining acceptable models. Norms of experimentation can be associated with the operationalisation of theoretical terms. They are directed towards stability and reproducibility. Their application constitutes one or more additional representational layers, which map theory-laden contexts to experimental situations (see Fig. 3.1).

There is no place here to compare this account of representation to other accounts that have been proposed in the literature. However, I hope that this presentation makes clear that the two-stage account is able to combine the contextual aspects

¹⁶One could object that perception can fail to represent (and not merely misrepresent) in the case of hallucinations, so that there is some form of licensing after all, and that for this reason, the fact that a mental representation is a representation of its target is not a pragmatic tautology at the communal level. One could respond that hallucinations count as representation of a non-existent target, which is different from not representing at all. Obviously, a lot more could be said on this topic, in particular on the accuracy of perception and on the way cognition and intentions shape perception, but all this is far beyond the scope of this book.

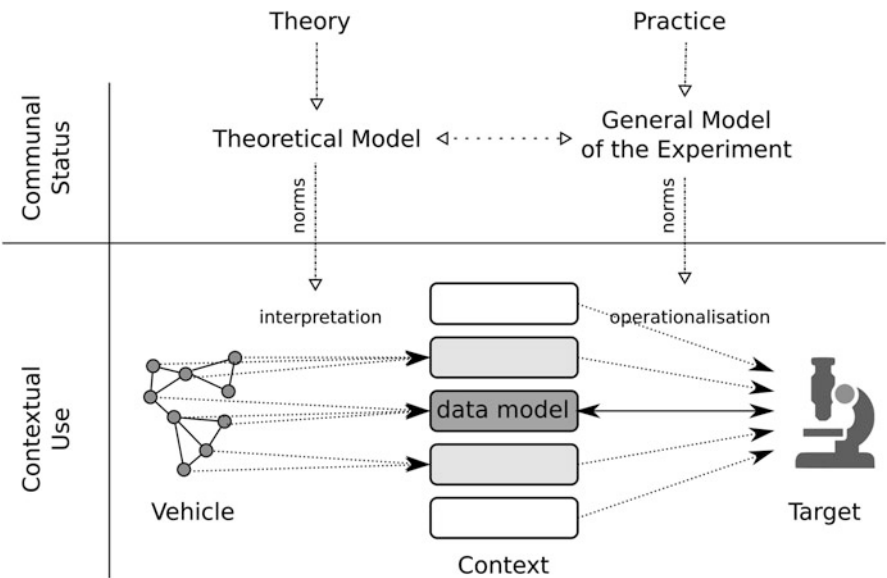


Fig. 3.1 The two-stage account of scientific representation

of experimentation and the unifying aspect of general models and theories. It accounts for most of the features of scientific representation and theories examined in Chap. 2, such as the sensitivity to purposes, the directionality of representation and the possibility of misrepresentation, the practical aspects of experimentation and communal licensing. The way it does so is by understanding unifying aspects in terms of norms constraining particular uses.

In the case of science, these norms are of four types. The three first types are distinctively theoretical, while the fourth type is relatively independent from theories:

1. the meta-norms conveyed by theoretical laws and principles, which constrain the legitimate models of a theory,
2. the compositional norms conveyed by abstract models, which constrain the construction of complex models from simple ones,
3. the domain-specific norms conveyed by indexical theoretical models, which constrain the way of representing particular types of phenomena, and
4. the experimental norms conveyed by models of experiments, which have to do with licensed experimental practices and operationalisation of theoretical terms.

The pragmatic nature of this account is reflected in the fact that the general representational status of models and theories is explicated in relation to potential contextual uses, and not the other way around: this is what I have called, in the introduction to this chapter, a practice-first approach.

By way of conclusion, let us now go back to the axiological debate that occupies us in this work: how shall we frame this debate in relation to the two-stage account?

Theories evolve. They are applied to new domains of experience, and sometimes they get replaced by new ones. All this can be characterised in terms of a selection of norms, and in particular, of meta-norms and domain-specific norms. The axiological debate concerns the ultimate finality of scientists when they construct and revise these norms: what are they after? What would be an ideal theory?

Norms and meta-norms can be evaluated according to various criteria. Norms can be simple or complex, many or few, specific or general. And more importantly in the context of epistemic representation, they can lead to successful or unsuccessful uses: what I have called accuracy. There are good reasons to think that empirical success is of primary importance for scientists, and that beyond the unificatory power of theories, a systematic connection to experience is one of the main characteristics that distinguish science from other human activities.

In the context of the account of representation presented in this chapter, the axiological debate about the aim of science can be understood as concerned with whether the values involved in the evaluation of norms of representation, and accuracy in particular, are instrumental for achieving a greater aim such as truth, or if they are sought for their own sakes. According to the realist, empirical success and other values, such as simplicity, coherence and scope, are indicators of truth, and science is after truth. Arguments concerning the truth-conduciveness of these values can be given to support this position (this is how the epistemic debate informs the axiological debate): these arguments will be examined in Chap. 6. In contrast, according to the empiricist, a certain notion of ideal empirical success at the theory level is the principal motivation for science. This notion, empirical adequacy, is enough to account for scientific practice, or the “rules of the game” of science, and there is no reason to assume that empirical adequacy is a means to achieve a greater aim.

A correct understanding of the notion of empirical adequacy, taken to be the main criteria by which representational norms are evaluated in science, is crucial for assessing the prospect of empiricism with regard to scientific realism. The present account of scientific representation is a useful framework for this purpose. The empirical adequacy of a theory must supervene in one way or another on the accuracy of its uses, since accuracy is a measure of empirical success. However, the right way of articulating model accuracy and theory adequacy is not straightforward. The reason for this is that accuracy is a contextual notion: it concerns one particular model interpreted in terms of one particular context, while a theory is constituted of more than one model, and each of them is applicable in more than one context.

The purpose of the next chapter is to make explicit this articulation between model accuracy and empirical adequacy by providing an analysis of empirical success for theories within the present two-stage account of representation, starting from contextual uses up to theories as wholes. This will result in a definition of empirical adequacy for theories in terms of the *relevance* and *accuracy* functions introduced in this chapter. Thus, the basis for a philosophical analysis of the aims and achievements of science will be laid out. And as we will see, this

definition of empirical adequacy will differ substantially from traditional empiricist understandings of the notion, in particular because of its modal character.

References

- Bailer-Jones, D. (2013). *Scientific models in philosophy of science* (vol. 43). Pittsburgh: University of Pittsburgh Press.
- Boesch, B. (2017). There is a special problem of scientific representation. *Philosophy of Science*, 84(5), 970–981.
- Bueno, O. (1997). Empirical adequacy: A partial structures approach. *Studies in History and Philosophy of Science Part A*, 28(4), 585–610.
- Callender, C., & Cohen, J. (2006). There is no special problem about scientific representation. *Theoria: Revista de Teoría, Historia y Fundamentos de la Ciencia* 21(1):67–85.
- Contessa, G. (2007). Scientific representation, interpretation, and surrogative reasoning. *Philosophy of Science*, 74(1), 48–68.
- Franklin, A., Anderson, M., Brock, D., Coleman, S., Downing, J., Gruvander, A., et al. (1989). Can a theory-laden observation test the theory? *British Journal for the Philosophy of Science*, 40(2), 229–231.
- French, S. (2017). Identity conditions, idealisations and isomorphisms: A defence of the semantic approach. *Synthese*, s11229–017–1564–z. <https://doi.org/10.1007/s11229-017-1564-z>
- Giere, R. (1999). Science without laws. In *Science and its conceptual foundations*. Chicago: University of Chicago Press.
- Grice, P. (1989). *Studies in the way of words*. Cambridge: Harvard University Press.
- Griffiths, R. (2003). *Consistent quantum theory*. Cambridge: Cambridge University Press.
- Hempel, C., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, 15(2), 135–175.
- Kaplan, D. (1989). Demonstratives: An essay on the semantics, logic, metaphysics and epistemology of demonstratives and other indexicals. In J. Almog, J. Perry, & H. Wettstein (Eds.), *Themes from kaplan* (pp. 481–563). Oxford: Oxford University Press.
- Kitcher, P., & Salmon, W. C. (Eds.). (1989). *Scientific explanation. Minnesota studies in the philosophy of science* (vol. 13). Minneapolis: University of Minnesota Press.
- Knuuttila, T., & Boon, M. (2011). How do models give us knowledge? The case of carnot’s ideal heat engine. *European Journal for Philosophy of Science*, 1(3), 309–334. <https://doi.org/10.1007/s13194-011-0029-3>
- Lakatos, I. (1978). *The methodology of scientific research programmes*. Cambridge: Cambridge University Press.
- Latour, B. (1993). Le topofil de Boa Vista ou la référence scientifique. In B. Conein, N. Dodier, & L. Thévenot (Eds.), *Les Objets dans l’action: de la maison au laboratoire*. Raisons pratiques, Editions de l’Ecole des hautes études de sciences sociales, Paris (vol. 4), 187–216.
- Mitchell, S. D., & Gronenborn, A. M. (2015). After fifty years, why are protein X-ray crystallographers still in business? *British Journal for the Philosophy of Science* 68, axv051. <https://doi.org/10.1093/bjps/axv051>
- Morgan, M., & Morrison, M. (1999). *Models as mediators: Perspectives on natural and social science*. Cambridge: Cambridge University Press.
- Nguyen, J. (2016). On the pragmatic equivalence between representing data and phenomena. *Philosophy of Science*, 83(2), 171–191.
- Ruyant, Q. (2021). True griceanism: Filling the gaps in callender and Cohen’s account of scientific representation. *Philosophy of Science* (forthcoming). <https://doi.org/10.1086/712882>

- Saatsi, J., & Pexton, M. (2013). Reassessing woodward's account of explanation: Regularities, counterfactuals, and noncausal explanations. *Philosophy of Science*, 80(5), 613–624. <https://doi.org/10.1086/673899>
- Salmon, W. (1984). *Scientific explanation and the causal structure of the world*. Princeton: Princeton University Press.
- Suppes, P. (1969). Models of data. In *Studies in the methodology and foundations of science* (pp. 24–35). Dordrecht: Springer. https://doi.org/10.1007/978-94-017-3173-7_2
- Suárez, M. (2004). An inferential conception of scientific representation. *Philosophy of Science*, 71(5), 767–779.
- van Fraassen, B. (1980). *The scientific image*. Oxford: Oxford University Press.
- van Fraassen, B. (1989). *Laws and symmetry* (vol. 102). Oxford: Oxford University Press.
- van Fraassen, B. (2008). *Scientific representation: Paradoxes of perspective* (vol. 70). Oxford: Oxford University Press.
- Vorms, M. (2011). Representing with imaginary models: Formats matter. *Studies in History and Philosophy of Science Part A*, 42(2), 287–295. <https://doi.org/10.1016/j.shpsa.2010.11.036>
- Weisberg, M. (2013). *Simulation and similarity: Using models to understand the world*. Oxford: Oxford University Press.
- Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford studies in philosophy of science. New York: Oxford University Press.
- Zahar, E. (1973). Why did Einstein's programme supersede Lorentz's? (II). *The British Journal for the Philosophy of Science*, 24(3), 223–262. <https://doi.org/10.1093/bjps/24.3.223>

Chapter 4

Modal Empirical Adequacy



Abstract Empirical adequacy, taken to be an ideal aim for theories, is undoubtedly an important notion for understanding scientific practice, and indeed the most important notion for an empiricist. This chapter examines empirical adequacy from the perspective of the account of scientific representation developed in the previous chapter. I adopt a bottom-up approach, which consists in two ampliative moves: first from model accuracy in particular contexts to model adequacy in general, then from model adequacy to empirical adequacy at the theory level. The resulting conception of empirical adequacy states that a theory is empirically adequate if it would be successful in all its possible relevant applications, that is, for all possible manipulations and observations we could make. This notion differs from van Fraassen's original proposal in three important respects: it is situation-based, pragmatic and modal. I argue that it makes better sense of scientific practice overall.

4.1 Empirical Adequacy as an Axiological Notion

As explained in the introduction of this book, the debate on scientific realism can be framed as an axiological debate concerning the *aim of science* (or perhaps the most important aim of science). This aim is not to be identified with the motives of individual scientists, which can be diverse, but with the goal of the community or institution of science, or the “rules of the game” of scientific practice. In particular, this aim determines what counts as the ideal success for a scientific theory, and what the criteria for their acceptance are.

Empirical adequacy is a notion introduced by van Fraassen (1980) to describe what he regards as the aim of science: producing theories that are empirically adequate. This notion is meant to account for the undeniable central aspect of empirical confrontation in science. According to van Fraassen's constructive empiricism, “[s]cience aims to give us theories which are empirically adequate; and acceptance of a theory involves as belief only that it is empirically adequate” (van Fraassen 1980, p. 12). In contrast, for a scientific realist, “[s]cience aims to give us, in its theories, a literally true story of what the world is like; and acceptance of a scientific theory involves the belief that it is true” (p. 8).

Van Fraassen defines empirical adequacy as follows:

[A theory is empirically adequate exactly if] such a theory has at least one model that all the actual phenomena fit inside. I must emphasize that this refers to all the phenomena; these are not exhausted by those actually observed, nor even by those observed at some time, whether past, present, or future. (van Fraassen 1980, p. 12)

What van Fraassen means by “fit inside” here is that some parts of a theoretical model, which he calls its “empirical sub-structures”, are isomorphic to data models that would represent all observable phenomena¹ in the universe (van Fraassen 1980, p. 64) (see also van Fraassen 1989, p. 228). The general idea is that all that is required to accept a theory is that it “saves the observable phenomena”. According to constructive empiricism, even if some individual scientists are more optimistic, and assume that the theory gives us “a literally true story of what the world is like”, empirical adequacy is still the main criterion by which theories are to be accepted or rejected in science. Note two interesting features of van Fraassen’s definition of empirical adequacy: it involves a reference to all phenomena of the universe, not only the ones that are actually observed, and it refers to a single model of the theory that could account for all these phenomena.

The aim of this book is to present and defend an empiricist position, which shares with constructive empiricism the idea that science aims to give us theories that are empirically adequate. The way this position differs from van Fraassen’s constructive empiricism lies entirely in the way it understands empirical adequacy. I am convinced that a notion of empirical adequacy that is more pragmatically oriented than van Fraassen’s is better apt to capture the aim of science, and I will argue that one of the main consequences of adopting such a pragmatic stance is that we have to assume that empirical adequacy is a modal notion, in a rather strong sense. This results in the position that I call modal empiricism.

The previous chapter was dedicated to providing tools for the construction of this position. This chapter presents the position itself, and its differences with constructive empiricism. This takes the form of a proposal for a definition of empirical adequacy, which derives quite directly from the account of epistemic representation presented in the previous chapter. Note that the idea that empirical adequacy constitutes the principal aim of science is not discussed here, but simply assumed. The rest of this book is dedicated to the defence of the resulting position on various grounds. However, the definition of empirical adequacy proposed in this chapter could be of interest even for a realist assuming that science aims at more than empirical adequacy.

Rather than criticising van Fraassen’s definition, what I propose is first to develop a different proposal “from scratch”, by reflecting on the way empirical results bear on the evaluation of the acceptability of a theory. This development phase is

¹For van Fraassen, being directly observable is part of the definition of phenomena (van Fraassen 2008, p. 8). This is faithful to the history of the concept, but the term is often extended to what can be measured with instruments. I will not follow van Fraassen’s terminology here, and talk about observable phenomena when they are directly observable.

based on the framework presented in the previous chapter, and in particular, on the *accuracy* function, which captures a notion of empirical success for models in particular contexts, and the *relevance* function, which expresses communal norms concerning the appropriate way of applying theoretical models in particular contexts. Remember that *relevance* determines conditions of accuracy for any context, and that *accuracy* tells us whether these conditions are fulfilled by an actual target of representation in context. The notion of empirical adequacy tells us what conditions the models of a theory, with their structures and *relevance* functions, should fulfill for the theory to be ideally successful. This requires reflecting on two aspects: first, what it would take for a model to be not only accurate in particular contexts, but ideally successful in all contexts, or *adequate*, and then, what the significance of the adequacy of a model is for the adequacy of the whole theory of which it is a model: how does the success or failure of a model affect the ideal success or failure of the theory? These two aspects are addressed in turn in the development phase.

After this development phase, I will compare the resulting definition of empirical adequacy to van Fraassen's original definition presented above. My definition differs from van Fraassen's in three respects: (i) it is expressed by quantifying over contexts and models, rather than by making reference to a unique model of the theory (ii) it is expressed in terms of the pragmatic notions of relevance and accuracy instead of the objective notion of observable phenomena, and (iii) it is explicitly modal, taking empirical adequacy to be about possible interventions and observations on (actual) situations. The last difference is probably the most radical departure from van Fraassen's notion, and it is the main defining characteristic of modal empiricism. However, as I argue in the last section of this chapter, these three aspects are related and make better sense of scientific practice as a rational activity overall. This definition is the best way to capture what scientists are aiming at as a community.

Note that van Fraassen (2008) has updated his views in several respects since the 1980s, notably emphasising the intentional and situated aspects of scientific representation. He has elaborated on the distinction between phenomena (observable entities, including measuring apparatus) and appearances (measurement results, data models, what phenomena "look like" from a given perspective). If anything, this should show the proximity of his views with the pragmatist stance adopted in this work. The reader could wonder why I focus here on a conception of empirical adequacy that was proposed 40 years ago, before these more recent developments. One reason is that the definition of empirical adequacy stated above is not explicitly updated in van Fraassen's recent work, even though a passage from van Fraassen (1980) using this definition is cited in van Fraassen (2008, p. 317) (see also p. 250). To my knowledge, no other proposals have been made in the literature, apart from discussions concerning the mathematical relation between data models and empirical sub-structures (Bueno 1997; Suárez 2005). This still seems to be the general understanding of empirical adequacy today. The other reason is that despite the pragmatist stance of his recent work, some important differences remain between van Fraassen's views and my own, and these differences do play an important role for our respective understandings of empirical adequacy. They

concern in particular the role of measuring instruments in experimentation and his reliance on observable phenomena. The views on scientific representation developed in van Fraassen (2008) do not necessarily call for an update of his old definition of empirical adequacy, which could explain why the old definition is not updated in this work, while my views do require a different definition.

4.2 From Contextual Accuracy to General Adequacy

Let us first examine what it would take for a model to be ideally successful, or *adequate*, in a way that is not restricted to particular contexts.

4.2.1 An Ampliative Notion

In Chap. 3, I have introduced a notion of *accuracy* that applies to interpreted models in particular contexts. Remember that an interpreted model is, roughly speaking, a symbolic structure such that the denotation of symbols is specified in terms of accessible properties of interest of the target. These properties of interest constitute the context. In effect, the interpreted model determines a set of “permitted” coarse-grained states or histories for the target, and it is accurate if the actual state or history of the target, as described by the data model, is among these possibilities. This notion of accuracy is not very different from van Fraassen’s idea that the empirical sub-structures of the model are isomorphic to data models, except perhaps for the fact that what counts as an empirical sub-structure and the way these sub-structures are compared to data models is determined by the context and purposes of the agent (something that is a priori compatible with van Fraassen’s more recent work).

Accuracy is a contextual notion that expresses actual empirical success. In contrast, empirical adequacy is an ampliative concept. It is a notion of ideal success that is not restricted to actual success. This is explicit in van Fraassen’s definition: empirical adequacy concerns not only actually observed phenomena but all observable phenomena in the universe, past, present and future, including unobserved ones. The reason for this is that the purpose of the notion is to account for the motivation of scientists as a community (what science aims at), and not solely for what they actually do, did or will do. This is an axiological component. So, this notion should do more than registering actual successes. The question that will occupy us in this section is how to understand this ampliative move, and in particular, how much extended it should be. It is at this level that the main differences between modal empiricism and van Fraassen’s constructive empiricism will emerge.

I take the notion of empirical adequacy to apply not only at the level of theories but also at the level of theoretical models. Theoretical models will be the unit of analysis in this section. Bear in mind that in the account of scientific representation

presented in Chap. 3, a theoretical model was defined as a symbolic structure associated with a relevance function that specifies appropriate or “licensed” interpretations of this structure for any context of use (in case no interpretation is licensed, the context is outside of the domain of application of the model). The model thus conveys domain-specific knowledge and postulates, as well as acceptable idealisations. We could expect a theoretical model to lead its users to empirical success in any context, for any of its licensed interpretations, in which case we could say that the model is empirically adequate. Ideally, this would not only concern past uses of the model, but any potential use. In order to capture this aspect, we need to talk about merely possible representational activities.

In van Fraassen’s definition of empirical adequacy, this extension to potential representational uses is captured by the “able” in “observable”. According to van Fraassen, the notion of being observable implies modal statements such as “In some circumstances, we would observe X” (van Fraassen 1980, p.16). However, van Fraassen is a modal sceptic. He would deny that the notion of being observable is modal and that such modal statements have objective truth-values, but he would claim that the property of being observable is objective nonetheless. This has generated controversy (Rosen 1994; Ladyman 2000; Monton and van Fraassen 2003; Ladyman 2004; Dicken and Lipton 2006). I will return to this controversy later on, but until then, let me just employ modal talk innocently and see where we can get to.

4.2.2 *Situations and Possible Contexts*

So, how shall we extend our account of empirical success to non-actual representational uses?

One important aspect of the representation relation is the denotational aspect: the user takes various parts of a model to denote various aspects of a target. In the framework presented in the previous chapter, this is captured by the notions of context and interpretation (Sect. 3.2). The context denotes various properties of interest, and the interpretation maps the model to the context. In order to talk about merely potential denotation, we will have to talk about merely possible contexts and interpretations. But we do not necessarily want the targets of denotation to be fictitious: possible contexts can be anchored to real objects.

In order to account for this aspect, let me introduce the notion of a situation. A situation is a local state of affairs, and a potential target of representation.² I assume that situations are occurrent and bounded in space and time, and that they can be delimited more or less arbitrarily. For example, we could consider the trajectory of a single planet, or of all planets in the solar system, over a few days or a few years: all of these are different situations. But once their boundaries are delimited, situations

²The notion can be compared to situation semantics in philosophy of language (Kratzer 2019).

have objective characteristics that are in principle accessible to us, for example, the relative positions of these planets. These are the characteristics that are typically involved in contexts and model interpretations. Typical examples of situations are the experimental situations from which experimental data are extracted, but they need not be, since a situation need not be actually represented. For example, a stone falling from a cliff on a distant planet counts as a situation because it *could* be represented (van Fraassen would say that it is *observable*). This counterfactual does not imply that the situation is not actual: only the fact that we represent it is counterfactual.

Let us assume that a situation can be associated with a set of possible contexts, and therefore, a set of possible ways it could be represented. Remember that in the framework presented in Chap. 3, a context captures the interests of agents, that is, the set of (theory-laden) properties from which they identify a relevant target, and the set of properties that they wish to inquire about. The idea of a possible context is that for any situation, agents could be interested in certain accessible properties of the situation or in others. I assume that we could enumerate these possibilities a priori: take any accessible object of the situation and any accessible quantity and imagine that they were relevant or not, with various standards of precision. Assuming the distinction between fixed and variable properties of interest introduced in Sect. 3.2.1, we could also imagine that this quantity would be known by the agent, or that the agent would inquire about it.

Note that as explained in Sect. 3.2.1, contexts are theory-laden entities (a user expresses her interests in theoretical terms). This means that the range of conceivable contexts depends on the situation, but also presumably on a theoretical perspective associated with a classification of relevant patterns in the accessible characteristics of the situation. One could say that the range of contexts afforded by a particular situation in the world is determined by the norms of experimentation which specify the empirical interpretation of theoretical terms in terms of those patterns (Sect. 3.4.2). These norms tell us what theoretical properties could be attributed to the target system in a particular situation, what state-space could be assigned to this system (in physics), and how the regions of this state-space can be associated with empirically accessible characteristics of the situation. They classify the phenomena in a certain way, so as to guarantee experimental stability. The interests of hypothetical users specified by a possible context are expressed directly in theoretical terms, so norms of experimentation determine the range of conceivable contexts.

Although relatively autonomous from theories, these norms are not entirely independent from theoretical frameworks. In physics, the structure of state-spaces evolves with theories. For example, measuring with great accuracy both the position and the velocity of an object might have been conceivable before quantum theory, but as we know, quantum theory imposes limitations on this possibility. So, some representational contexts that were considered a priori possible a few centuries ago (being interested in both position and velocity, made accessible with unlimited precisions) are not considered possible today, because our understanding of these theoretical properties has evolved.

However, this is not necessarily problematic. The fact that the range of contexts afforded by a situation depends on experimental norms, which in turn are affected by theories, will not entail any vicious circularity between our theoretical commitments and the empirical adequacy of our theories, nor any kind of “experimental regress”. It implies, at most, a partial perspective on reality. Merely defining a range of representational contexts available for a given situation is insufficient to claim that our models are accurate in these contexts. Furthermore, experimental norms are still relatively autonomous from theories. And finally, it seems plausible that the limitations brought by quantum mechanics are empirically grounded, so that a theory interpreted by means of experimental norms that do not incorporate these limitations on position and velocity could not be empirically adequate anyway (perhaps, for example, these norms could not sustain the phenomenal stability of position and velocity measurements, because of problematic operationalisations of these quantities: see Sect. 2.3.3 on the role of stability).

If, assuming experimental norms, the possible contexts afforded by a given situation can be enumerated a priori in the way specified above, by listing properties of interest, then the relevant modality for contexts seems to be something like logical possibility or conceivability (within the theoretical framework), and I will talk about conceivable contexts from now on. I should note that no inference from conceivability to metaphysical possibility is assumed here, and also that the notion of conceivability involved is more restricted than the one sometimes entertained by metaphysicians, since for the moment, we are not conceiving alternative states of affairs, let alone other possible worlds, but only alternative perspectives on a bounded and actual state of affairs. For any given set of experimental norms, the set of conceivable perspectives entirely supervenes on this state of affairs, and these perspectives only concern empirically accessible characteristics.

Let us introduce a new function, which specifies the set of conceivable contexts afforded by a given situation.

Definition 4.1 *affordance*(S): specifies the set of conceivable contexts $\{C\}$ afforded by the situation S .

This function could have taken a set of experimental norms as a parameter, but I will omit it for the sake of simplicity.

We can say that a situation, viewed from the perspective of a particular context (a set of relevant coarse-grained properties), is in a certain state. This state is, intuitively speaking, the projection of this situation onto the context: it is restricted to the properties of interest specified by the context, interpreted by means of experimental norms, given some expected standards of precision. An interpreted model is adequate if this state is among the ones that are “permitted” by the model (see Sect. 3.2.4).

The *relevance* function introduced in the previous chapter specifies which interpretations of a theoretical model are licensed by the epistemic community in which contexts. Let us assume that the *relevance* function can take a conceivable context as a parameter. This means that the norms that guide the application of a theoretical model can be extended to counterfactual representational uses. For

example, there is a fact of the matter about how a Newtonian model of free fall would apply to the fall of a stone on a distant planet if agents were interested in the trajectory of this stone with a given degree of precision (assuming scientists know the mass of this planet), even if this model is never actually used in such a context. This assumption that *relevance* applies to merely conceivable contexts is required for the ampliative move we are interested in: if there were no fact of the matter as to which interpretations are counterfactually legitimate, we could only talk about accuracy for actual representational activities, and the notion of empirical adequacy would only concern actually observed phenomena, not all observable phenomena. This could function as a description of what scientists do, but not as an account of their aspirations.

4.2.3 Different Versions of Empirical Adequacy

What I propose, in order to define empirical adequacy for theoretical models, is to quantify over all situations in the universe, and over all the conceivable contexts associated with these situations. This should be taken as a mere transposition, in our account of representation, of van Fraassen's locution "all observable phenomena in the universe". Empirical adequacy is then defined in terms of the accuracy of the licensed interpretations of a model in all these conceivable contexts. A theoretical model is empirically adequate if for any situation in the world in which it would have licensed interpretations, the model would be accurate if it were used to represent it.

We could wonder whether the quantification over contexts and interpretations, for a given situation, should be existential or universal. Maybe it would be enough that the model is accurate for *at least one* conceivable context or interpretation. But arguably, universal quantification makes more sense. Remember that contexts of representation specify standards of accuracy and objects and properties of interest. They determine acceptable idealisations. I assume that in general, we do want empirical success whatever the kind of properties we are interested in, and ideally, for all standards of accuracy for which the model has licensed interpretations. If it were enough to have success for at least one level of idealisation in a given situation, empirical adequacy would be quite easy to achieve: we could just lower expected levels of precision for the predictions of the model, or let the set of salient properties involved in the interpretation be nearly empty.

We can then assume the following definition for adequacy, where S refers to actual situations in the world:

Definition 4.2

$$adequacy_a(M) := (\forall S)(\forall C \in affordance(S))(\forall m \in relevance(M, C))accuracy(m)$$

A theoretical model is adequate if for any situation and any representational context afforded by this situation, the licensed interpretations of this model in this context are accurate.

We already notice a slight difference between this definition of empirical adequacy and van Fraassen's. It lies in the fact that a model is assumed to be interpreted contextually, while van Fraassen's definition apparently assumes something like a cosmic interpretation in terms of all observable phenomena in the universe. Van Fraassen's definition functions as if there was a cosmic context associated with a universal situation in which models could be interpreted, assuming a univocal way of comparing data models and the empirical sub-structures of the model. This aspect is in tension with his more recent writing on scientific representation, and perhaps he would not object to this alternative formulation. I will discuss this difference in Sect. 4.4.1.

Apart from this difference, the definition is quite compatible with constructive empiricism. In particular, it extends empirical success to all observable phenomena, but without considering merely possible phenomena: we are still quantifying over actual situations. But note that it could easily be adapted to other flavours of empiricism by changing the domains of quantification: what if instead of all actual situations, we only quantified over situations and contexts on which we have experimented so far or will experiment? And what if, instead, we quantified over all possible situations? This would give us the following versions of empiricism:³

- Manifest empiricism: Empirical adequacy concerns all actual situations and contexts on which we have experimented or will experiment.
- Actual empiricism: Empirical adequacy concerns all actual situations on which we could *conceivably* experiment.
- Modal empiricism: Empirical adequacy concerns all *possible* situations on which we could conceivably experiment, even those that are not actual.

We have already given a reason to reject manifest empiricism: an extension to merely potential representational uses is required for empirical adequacy to be axiological (I will return to this reason, which is endorsed by van Fraassen, in Sect. 4.4.3). Van Fraassen's constructive empiricism is a version of actual empiricism. The question that I wish to ask now is: why not modal empiricism? Could not it be a better way of accounting for scientific practice?

There are good reasons to think that we *must* adopt a modal empiricism rather than an actual one, as I will now explain.

³Giere has proposed a similar taxonomy with slightly different labels in Churchland and Hooker (1985, ch. 4).

4.2.4 *Are Modalities Innocuous?*

There is an implicit modal component in the definition of empirical adequacy given above, which lies in the quantification over conceivable contexts. In so far as situations are taken to be actual states of affairs, this modal component is the counterpart of van Fraassen's "able" in "observable". In order to argue for modal empiricism, let us examine this modal aspect. Can we make sense of it in an innocuous way, that is, without assuming that there are real possibilities in the world? In particular, what are the implications of this quantification over possible contexts with regard to the *accuracy* function defined in the previous chapter? Can actual states of affairs act as truth-makers for this function, even for counterfactual contexts?

It seems so, at first sight. The *accuracy* function specifies whether the actual state of a target is "permitted" by an interpreted model. We can conceive various representational contexts where a general model would be interpreted in a particular, licensed way in terms of a given situation, even if these interpretations do not correspond to any actual representational use. This only means that no scientist is actually using the model in these ways to make inferences on this situation. We have assumed that these conceivable contexts can be enumerated a priori, and this process of imagination does not seem to affect the situation itself, but only the way it would be represented. We could think that the actual state or history that could be assigned to the situation from various contexts would supervene on a real, underlying state. So, there seem to be an actual fact of the matter about whether the actual state of this situation is permitted by such interpreted models, even if these models are not used. For this reason, it could seem that the implications of an extension to counterfactual uses are not very important. The modal component involved seems innocuous.

But not so fast. I claimed in the previous chapters that the direction of fit when interpreting a model is not necessarily from the model to the target. In many contexts, interpreting a model in terms of a target involves an adjustment between the target, the model and the contextual environment: it is not only about finding the right model and interpretation for the target, but also about preparing the target in the right way, calibrating our instruments, positioning them correctly and so on. In other words, experimentation is not a matter of passively observing the world, but also a matter of intervening appropriately. As observed in Sect. 3.4.2, in science, the context of use will generally be a model of the experiment, which implies the presence of measuring apparatus and their manipulation. This could be incompatible with the actual state of the situations being considered. Shall we say that such counterfactual contexts are still afforded by the situation?

Take as an illustration the manipulations involved in "observing" proteins. We have molecular models of proteins, but to confirm the accuracy of these models, one should break down the cell membranes, for example using enzymes, purify the resulting mixture by centrifugation and isolate the proteins of interest by using various types of chromatography on the basis of their known chemical characteristics; only then can their properties be observed, for example through

spectroscopy.⁴ A molecular model of a protein can certainly not account for our direct observations of the behaviour of living organisms; it can only account for the result of particular intentional operations. So, a situation involving a living organism in its natural habitat does not directly afford contexts where a model of protein can be interpreted. Nevertheless, there seems to be a connection between the model and the situation (between models of proteins and living organisms): what could it be? An option is, of course, to go realist and claim that the connection is that some unobservable entities represented by the model are present in the experimental situation, and also in living organisms. But what are the empiricist alternatives?

We observed in the previous chapter (Sect. 3.3.5) that models can be combined into more complex ones, taking the example (which I owed to Mitchell and Gronenborn) of the model of a protein integrated into a model of virus propagation. In general, what will be relevant for representing a living organism in normal conditions will not be a simple protein model, but a more complex model that integrates this protein model. So, perhaps the protein model is not directly relevant for living organisms after all, but only indirectly relevant in so far as it can be used to construct more complex models. Although this is an important observation, it does not entirely solve our problem, because the progression of a virus in an organism also needs to be monitored by appropriate interventions, such as taking a sample and analysing it. However, I presume that we consider that a model representing the progression of a virus in an organism is *potentially* relevant for any infected living organisms, including when this progression is not monitored. So, again, we need a connection between the situation where the model can be interpreted directly in terms of accessible characteristics and other situations where this is not the case.

In order to address this question, let us first consider another question, which concerns representation in situations where the appropriate manipulations are performed, for example, situations in laboratories where proteins are observed, or situations where the virus is monitored by taking samples. Van Fraassen (2008, pp. 99–100) would say that in such situations, scientific instruments “enlarge the observable world” by creating new observable phenomena. But how shall we interpret this idea? In these contexts, shall we say that interventions are part of the context of observation, or part of the observed situation?

Here, we face a dilemma. If we want to say that interventions are observed, and that as such, they belong to the target of representation, then a model of a protein is only relevant for the particular type of situation where various manipulations are performed: by no means is it a model that represents a feature of living tissues in general, since none of the properties it denotes are accessible in most situations involving living tissues. And the same would go for the standard model of particles in physics, for example: its models would only represent highly controlled situations in colliders, certainly not characteristics of all material objects. It would

⁴There are actually different experimental protocols, and biological models of proteins result from an integration of multiple experimental models (Mitchell and Gronenborn 2015). This point will become relevant later.

be misleading to say that these models represent microscopic phenomena: colliders are very big, and paradoxically, the more “microscopic” the phenomena a theory is said to describe, the bigger the situations that it can represent are, in general.

All this seems quite wrong. In particular, it is hard to account for the motivations of scientists: why do they try to produce these new exotic phenomena, when they could be content having adequate models for the observable phenomena that we already have? Why spend so much money to create extremely weird, highly controlled unnatural situations such as colliders? Why develop models that only ever apply to these very particular cases? This looks like an unfaithful description of their activity.

On the other hand, we could claim that interventions are not part of what is observed: they merely bring it about. But if interventions are not really part of what is observed, then what is? Numbers on computer screens? The positions of needles? This does not look like a very interesting “enlargement of the observable world”.

The point is that interventions are not the main objects of our observations, but they are not mere instrumental means to produce new observations either: they are constitutive of the way scientific observations are interpreted. A number on a screen counts as (represents) a characteristic of a protein because various manipulations were intentionally performed beforehand, and because these manipulations are interpreted as a measurement of this characteristic.⁵ These interpretations are typically determined by norms of experimentation, for example, by models of experiments, and the particular operationalisation that is chosen is, in some respect, irrelevant in so far as it is known to be effective. The properties that are denoted by an interpreted model are not mere numbers on a screen: they are more accurately described as the results of such or such type of intervention.

So, the dilemma is the following: we need to take into account experimental interventions to correctly interpret the properties of the situation, but the focus of our interest is not on these interventions. We are interested in ordinary situations, not only the finely controlled ones that are found in laboratories.

The solution I propose to combine these two aspects is to go modal: we have to assume that the counterfactual contexts that interest us, the ones where particular interventions are performed, can correspond to contexts afforded by *alternative ways actual situations could be*. A pragmatist cannot avoid making our manipulations relevant characteristics of the situations being represented. However, it could be enough that these manipulations are possible, that is, that they characterise alternative ways a situation could be, in order to be relevant for a given situation. The truth-makers for the accuracy of counterfactual representational uses are counterfactual states of affairs: for example, what would happen if such or such manipulation were performed on an actual living tissue, or what would be observed if we monitored the progression of the virus by taking samples. However, these counterfactual states of affairs are anchored to actual situations. This way, our models can be about ordinary situations, and not about exotic ones, even when

⁵Van Fraassen (2008, ch. 6–7) makes similar remarks.

interpreting them implies taking into account complex interventions. The contexts where a model of virus progression would be licensed are not afforded by situations involving a living organism that is not monitored, for instance. But they could be, so the model is still relevantly connected to living organisms infected by the virus. This is an empiricist alternative to the realist view that the model does represent unobservable entities.

The main implication of this account is that the modal component of empirical adequacy, which is required to extend this notion to unobserved, but observable targets, cannot be innocuous. A context specifies not only levels of precision and salient objects, but also the kinds of operations that are performed on the target (what is measured, etc.). Perhaps possible operations on an actual target can be enumerated a priori, but the potential results of counterfactual operations cannot be known a priori, nor can they be identified with actually observed properties of targets unless these operations were performed. Such counterfactual contexts apply to alternative ways situations could be, so to mere possibilities, and the modality involved must be alethic (for example, nomological).

The potential results of counterfactual manipulations could be interpreted as manifestations of actual properties. But then these properties would be dispositions, not categorical properties, and we have a modal aspect in any case. The notion of disposition is metaphysically loaded. If, following an empiricist stance, we wish to remain metaphysically cautious, let us just talk about merely possible observations resulting from conceivable operations on an actual target.

4.2.5 *Exotic Possible Situations*

The upshot of these analyses is that an ideal desideratum for a good theoretical model is that whatever manipulations could be performed in a situation, including preparing the situation in a particular way and measuring particular quantities, the model would be accurate. More precisely, whatever the conceivable context afforded by any alternative way a situation could be,⁶ and whatever would eventually happen in this context, the interpretations of the model that would be licensed in this context would be accurate. This is what modal empiricism claims.

This commitment to natural possibilities, conceived of in terms of alternative ways actual situations could be, makes the position stronger than traditional versions of empiricism, but one could also fear that this way of thinking about natural modalities is too limited. Anchoring possibilities to actual situations is a way of avoiding the idea that theories and models would only concern exotic situations.

⁶This notion will be made more precise in the next chapter. For now, let us just say that these alternative ways the situation could be correspond to the same situation of the same type, with the same objects, but in a different state with regard to accessible properties, or with regard to the manipulations performed in this situation.

However, many times, scientists actually want their models to be about exotic possible situations with no actual counterparts. For example, we can consider possible technological developments based on new theories, or possible experimental designs that would confirm or disconfirm our theories, and these possibilities do not correspond, or are not necessarily known to correspond, to mere alternative ways of manipulating actual objects. Should an empiricist assume that such “exotic” situations are possible? If not, then should we be sceptical about the ability of our theories to account for them? And then, how shall we interpret scientific discourse?

First, note that the kind of conceivability by which such possibilities are entertained seems more metaphysically loaded than the one that has been entertained so far, with regard to the possible manipulations one could perform on an actual situation. Alternative ways actual situations could be involve different states for these situations, but I assume that they are still situations of the same type. Taking samples from an organism does not significantly alter this organism. On the other hand, considering a possible experimental design or a possible technology that has never been implemented means considering an entirely new type of situation. Since such possibilities are not directly anchored to actual situations, it is less clear whether they are possibilities *in our world* or *in another possible world*. This could bring suspicion from an empiricist perspective.

This question about the existence of possible situations detached from actual ones is related to the metaphysical debate between actualism and possibilism. For an actualist, only actual objects exist, and they can have modal properties, but merely possible objects do not exist, while for a possibilist, they do. Here, the relevant objects are situations, and what has been presented so far is a version of actualism with regard to situations: actual situations could have different properties than the ones they actually have, but merely possible situations unanchored to the actual world are not assumed to exist,⁷ so the possibilities considered are possibilities “in the actual world”. This actualist stance plays an important role in this work. As I will argue in the next chapter, it makes the kind of modality involved weaker than metaphysical or nomological modalities, and this is what primarily distinguishes modal empiricism from realist positions such as structural realism. However, the idea that an empirically adequate theory should ideally account for entirely new physical configurations will also play a role in my arguments against realism, in particular when it comes to accounting for novel, unexpected predictions, and I think that we should account for scientific discourse in this respect. There is a tension between actualism and this desideratum.

The solution I suggest is the following. I claimed earlier that situations can be delimited arbitrarily. In particular, they can be arbitrarily large. Assuming this, it is *prima facie* conceivable that many physical configurations that scientists could think of are possibly realised in some alternative to a large enough situation, past, present

⁷I will sometimes employ the locution “possible situations” to refer to possible ways actual situations could be in this book. This locution should not be interpreted as a commitment to possibilism.

or future. One could take, for example, an actual situation in a laboratory where all the instruments required to realise the configuration or technological design we are interested in are present. The alternative to this situation that interests us is one where the scientists of this laboratory, instead of doing what they are doing, perform all the actions necessary to realise the relevant configuration. It is *prima facie* plausible that (i) the alternative to the large situation that interests us is not only conceivable, but also naturally possible (that it is not precluded by natural constraints), and (ii) in so far as a model of the theory would, in principle, be apt to represent the large situation, and that it would be empirically adequate for this large situation and its possible alternatives, it would also correctly account for the configuration we are interested in, which is a sub-part of one alternative way this large situation could be.

In sum, my proposed solution consists in reinterpreting modal talk in science (and possibly elsewhere) in terms of sub-parts of alternative ways large enough actual situations could be.⁸ It is not necessary that we locate these possible situations precisely in so far as we assume that these possibilities exist.

I said that assumptions (i) and (ii) above are “*prima facie* plausible” because these assumptions are not certain, which means that a modal empiricist should be wary. Assumption (ii) is the less problematic of the two: it could be argued that if a complex model is empirically adequate, then the component parts of this model are also empirically adequate, at least as far as they can be interpreted in terms of accessible characteristics (this assumption could rest on the adequacy of compositional norms and on the coherence between experimental and theoretical norms).

Assumption (i) is more problematic. We do not have actual scientific models that can represent the very large situations involved in this line of reasoning. Such a model could inform us on whether (i) is true, assuming that our theory correctly accounts for what is possible or not, but it is, in general, unavailable. I presume that most of the time, the assumption that the realisation of an exotic physical configuration is possible rests on common sense: after all, we know a great deal about what is possible or not in this world. However, common sense is not always reliable, and our capacity to imagine all the details of possible alternatives to large situations is drastically limited in comparison with the way we can simply list the conceivable manipulations and outcomes of a concrete experiment. It might be the case that, unbeknownst to us, the large situations where the configuration we are interested in is realised are in fact naturally impossible.

This sceptical result is actually very welcome: indeed, as engineers and experimenters are well aware, implementing a new experimental or technological design is almost never an easy task. Many unexpected difficulties can appear along the way. A consequence of our interpretation of modal discourse in terms of alternatives to large enough situations is that the more distant from what has already been realised a

⁸I explore a potential formalisation of this combination of mereological and modal features in my PhD thesis (Ruyant 2017, ch. 9).

new physical configuration is, the larger the situation that must be considered (if new instruments must be manufactured, we should include this as well in the situation), and the more wary one should be that this type of situation is actually possible. This matches our common-sense intuitions regarding the fact that remote possibilities are less credible than proximal ones. Incidentally, this gives the modal empiricist good reasons to be radically sceptical about our ability to gain knowledge about purely metaphysical possibilities, in particular when they are expressed in terms of possible worlds: these are the most remote kinds of possibilities that we could ever conceive of.

To sum up, the modal empiricist believes that an empirically adequate theory can correctly account for all alternative ways actual situations could be. Considering large enough situations, this gives *prima facie* good reasons to assume that any physical configuration or any exotic type of situation we could think of could be accounted for by the theory. The modal empiricist needs not exclude such exotic situations from her account. However, the kind of assumption just mentioned is conditioned on the realisability of these exotic situations, and this realisability is not, in general, a direct consequence of the empirical adequacy of the theory, at least not one to which we have direct cognitive access, which brings sane scepticism.

4.2.6 Modal Empirical Adequacy

In light of the previous remarks, we shall retain an actualist understanding of modalities, in terms of possible ways actual situations could be. We can now propose the following definition for empirical adequacy:

Definition 4.3

$$\begin{aligned} &adequacy_m(M) := \\ &(\forall S)\Box_S(\forall C \in affordance(S))(\forall m \in relevance(M, C))accuracy(m) \end{aligned}$$

The only difference with the previous definition lies in the presence of the modal operator $\Box_S$. This operator quantifies over all possible ways a situation could be, and the notion of possibility involved is alethic: these are possibilities “in the world”, not “in the head”. In some of these possibilities, particular actions are performed, and they yield particular results. This means that particular conceivable representational contexts are afforded. They are given by the *affordance* function. The *relevance* function gives the set of interpretations of a general model that would be licensed in these contexts. The *accuracy* function returns true in case the state or history that would occur in this possible situation is permitted by an interpreted model m . The general model is empirically adequate if its interpretations that would be licensed in these possible situations would be accurate, whatever the possible situation.

More should be said about this kind of modality: one could wonder how to delimit the range of possible ways an actual situation could be, what role our theoretical perspective plays in delimiting this range, what the situation's trans-possibility identity conditions are, etc. All this will be addressed in Chap. 5.

For now, we can remain at an intuitive level and just assume that there are concrete situations in the world, alternative ways they could be, in particular assuming that manipulations could be performed on them, and natural constraints on these possibilities. This definition of empirical adequacy says, in substance, that a theoretical model is apt to capture the way these natural constraints affect the possible representational contexts in which this model applies. It is this aspect that, as I will argue in the following chapters, is able to respond to the main arguments in the debate on scientific realism, without assuming that our theories are true descriptions of a mind-independent reality.

Remember that adequacy is an axiological notion. It corresponds to criteria of acceptance by the scientific community. According to modal empiricism, a theoretical model is acceptable if scientists have good reasons to believe that this model is empirically adequate in the modal sense: the model could withstand any possible situation to which it would apply (the empirical reasons we have for believing this, that is, questions of justification, are also the subject of the next chapter). The model would be accurate "come what may", as long as it is correctly interpreted. This must be the case of the models currently accepted by scientists.

An important feature of this definition is that empirically adequate models allow for counterfactual reasoning. Take, for example, the model of a pendulum interpreted in such a way that the initial position of the pendulum is not fixed by the context. It makes sense to claim that if the initial position of the pendulum were x_0 , then the position of the pendulum at t would be x_1 , even when the initial position is not actually x_0 , so long as the model says so. This is because the same model interpreted in the same way could be used in the alternative situation where the initial position is x_0 , since this initial position is not fixed. Assuming that the theoretical model used in this context is modally adequate, this interpreted model would be accurate as well in this alternative situation.⁹ This constitutes an advantage for modal empiricism, since the position can directly account for parts of scientific discourse, for example, explanations and causal discourse (see Sect. 3.3.6).

We have made the most important step in our understanding of modal empiricism. However, modal empiricism is not a position about theoretical models, but a position about scientific theories, so we still need one more extension in order to reach our final definition. This will bring important reflections on the role played by relevance and on the unificatory power of theories.

⁹This reasoning assumes that there is an alternative possibility where the initial position is x_0 . A further requirement for counterfactual reasoning, beyond empirical adequacy, is for a model to be perfectly informative, in that it does not represent impossible states of affairs. This is briefly discussed below.

4.3 From Models to Theories

In the previous chapter, I have proposed that we understand theories as families of general models (Sect. 3.4.1). I noted that several options concerning the extension of this family are available. One could consider that all conceivable models satisfying the general laws and principles of the theory are models of the theory. Alternatively, one could consider that only some of them are: the ones that are licensed by the scientific community, because the domain-specific postulates that these models incorporate are fully part of the theory. In the first option, a theory does not evolve with time, while in the second one, it is enriched every time a new model is proposed to account for a new type of phenomenon.

We now have an understanding of when theoretical models can be considered ideally successful. The question that I wish to ask now is: how does the empirical success or failure of a model of a theory bear on the empirical success or failure of this theory? Answering this question is required in order to introduce a notion of empirical adequacy that applies to theories instead of models. However, as we will see, the answer depends on the way the family of models that constitute a theory is understood.

4.3.1 *Two Options*

As proposed in the previous section, we can say that a theoretical model is empirically adequate if all its licensed interpretations in any possible context are accurate. If one licensed use of a theoretical model leads to an empirical failure, then we know that the model is not adequate. But how should this discovery affect the theory of which it is a model?

The main problem that needs to be addressed in this respect is that an experimental failure does not necessarily mean that a theory should be abandoned, even if it concerns a model of the theory.

Let us illustrate this problem with the following scenario: imagine scientists want to model the solar system. They produce a Newtonian model. However, they discover that the model is not adequate: it does not account for the trajectory of Uranus. For this reason, they propose a new model with a yet unknown additional planet that they call Neptune, and calculate the position that this planet must have to account for the discrepancy in the old model. They point their telescope in this direction and discover that indeed, Neptune exists. Although an experimental failure was initially involved, this turned out to be a great empirical success for the theory: in a sense, the theory correctly predicted the existence of Neptune.

In other circumstances, experimental failures are indicative of theoretical failure. Consider a similar situation: scientists want to model the solar system, but their model does not match their observations with regard to the trajectory of Mercury. For this reason, they propose a new model with an additional planet, Vulcan.

Although the new model is accurate with regard to the trajectory of Mercury, Vulcan is never observed. A few decades later, the theory of general relativity is proposed, and it correctly accounts for the trajectory of Mercury without postulating any additional planet: retrospectively, the initial experimental failure in accounting for the trajectory of Mercury appears to be a failure for Newtonian mechanics as a whole.

These well-known historical examples, discussed by Kuhn (1962), show that the inadequacy of a model does not always entail the inadequacy of the associated theory. However, sometimes it does, or should. How shall we account for this?

There is room for interpretation, and it depends mainly on how we identify theories. Consider these two options, with regard to unproblematic experimental failures (the case of Neptune):

1. *A model without Neptune is relevant, but inadequate.* It is licensed as a legitimate representation of the solar system in Newtonian mechanics, but it makes incorrect predictions. Newtonian mechanics can still be considered an adequate theory, though, because another Newtonian model (one incorporating Neptune) is apt to represent the solar system adequately.
2. *A model without Neptune is irrelevant.* It should not be licensed by the community as a representation of the solar system, at least for contexts where enough precision with regard to the trajectory of Uranus is required. Newtonian mechanics is still an adequate theory, for the simple reason that irrelevant models do not jeopardise a theory's adequacy.

In the first option, norms of relevance are liberal. Perhaps only the correct interpretation of theoretical vocabulary (interpreting position as position) and the conformity with general laws and principles of the theory matter. This corresponds to the first way of identifying the family of models of the theory mentioned above. Because of this liberalism, many Newtonian models are apt to represent the solar system. Many of them are inadequate of course, but this does not matter: all that matters is that the theory is capable of providing at least one adequate model for any situation in its domain of application.¹⁰ The scientists' job, when extending the theory to new domains of application, is to find one adequate model, so as to confirm that it has at least one. In this respect, we could say that the discovery of Neptune extended the domain of application of the theory, with one more celestial object being a potential object of interest, and confirmed the adequacy of the new model with Neptune for this new extended domain.

In the second option, norms of relevance are strict. They include domain-specific knowledge. A model should get things right to be considered apt to represent a target system, and few Newtonian models meet this condition. This corresponds to the second way of identifying the family of models of a theory that I mentioned earlier. We can presume that it is not enough that at least one model be adequate: relevant

¹⁰This option seems to be the one implicitly adopted by van Fraassen: in his definition, a theory is empirically adequate if "at least one model" is such that all observable phenomena fit inside.

models should all be adequate, or we have a problem with the theory. The scientists' job, when extending the theory to new domains of application, is to find which models are relevant. Norms of relevance are revisable and sensitive to empirical inputs, and a theory can evolve following experimental failures. In this respect, the initial failure to predict the trajectory of Uranus cast doubt on the relevance of the initial model, and the discovery of Neptune confirmed these doubts, as well as the relevance of the new model with Neptune.

Which option is the right one? The question is whether unproblematic cases of experimental failure should count as inadequacy or as irrelevance, and as we can see, this question is strongly related to the problem of how to identify theories discussed in the previous chapter: shall we identify a theory with all models satisfying general laws and principles, in a liberal way, or shall we say that domain-specific norms are also part of the theory, thus bringing more constraints on its applicability?

This could look like a mere verbal dispute over the meaning of "relevant" and "adequate". However, I believe that this dispute is substantial, and that we can settle this case in favour of the second option: taking the model without Neptune to be irrelevant, and therefore unproblematic for Newtonian mechanics.

4.3.2 *Unproblematic Failure as Inaccuracy*

What could count against the second option, and in favour of the first one, is the following rationale. It seems quite plausible that in general, we learn that a model is not the right one by testing its adequacy. But if not being the right model just means being an irrelevant model, one that should not be part of the theory, then relevance collapses into adequacy, and the adequacy of the theory is achieved at a cheap price: instead of saying that a theory is inadequate, we exclude problematic cases from its domain of application, by saying that the corresponding models are irrelevant. Misrepresentation is impossible in this context. A theory cannot fail to be adequate.

This conclusion does not necessarily follow, because relevance can be sensitive to adequacy without collapsing into adequacy, as I will explain later. In any case, according to this rationale, we should be liberal and accept many interpreted models as relevant, perhaps all models satisfying the general laws and principles of the theory, so as to make room for misrepresentation. In general, many relevant models will be inadequate, and since we do not want to say that all theories are inadequate for this reason, we must assume that a theory can be adequate if only one of its relevant models is adequate for any situation in its domain of application. As noted in the previous chapter (Sect. 3.4.1), with this option, the mediation of theoretical models becomes superfluous, because any kind of domain-specific norm of relevance is a priori allowed as long as theoretical laws and principles are respected. We might just as well say that the theory must have at least one contextual interpretation that is accurate for any possible context in its domain of application, as if the theory was one very abstract model that could be interpreted in many ways.

However, this option has a serious flaw. It seems that if we are liberal enough with regard to conditions of relevance, it will be quite easy to find an adequate interpretation of a theory for any situation. In the case of Mercury presented above, the theory has an adequate model after all: the one with Vulcan. Just assume that this model is only relevant in situations limited to the trajectory of Mercury, and the model will be adequate. However, we would not say that Newtonian mechanics is adequate with regard to the trajectory of Mercury. In other words, having at least one accurate interpretation for any context might be too easy to achieve for a theory, and we could suspect that scientists have more stringent criteria of success.

One could respond to this objection that this difficulty disappears if we consider various situations and contexts: the theory should account for the trajectory of Mercury, but it should also tell us whether there is a planet between Mercury and the Sun. Newtonian mechanics seems to fail in this respect. But this is not true: Newtonian mechanics has at least one model that does not predict the presence of a massive body between Mercury and the Sun. In this option, no constraint is put on the coherence of models across various contexts, so we can take the model with Vulcan to account for the trajectory of Mercury, and the model without it to account for the absence of Vulcan.

Perhaps the problem can be solved by considering a larger situation that comprises both the trajectory of Mercury *and* the area between Mercury and the Sun. There Newtonian mechanics apparently fails, because a single model cannot account for both at the same time. As we can see, the option leads us to consider larger and larger situations, bringing more and more constraints on adequacy so as to avoid triviality, and we might soon be led to consider representing the universe as a whole, thus justifying van Fraassen's definition of empirical adequacy (remember his mention of a model such that *all phenomena in the universe* fit inside). We could entertain the idea that our theory, now construed as a single abstract model, has interpretations in a universal context where all accessible properties in the universe are relevant. The theory is empirically adequate if at least one such interpretation is accurate. The modality examined in the previous section would be naturally accounted for using the traditional possible world semantics. And perhaps assuming that such a model of the universe exists, other interpreted models become superfluous for adequacy, because they are all contained in this cosmic model. They would merely concern particular applications of the theory, but they would be irrelevant for expressing what the adequacy of the theory amounts to.

I will argue in Sect. 4.4.1 that such a move is very idealistic and brings us further from actual scientific practice, which is problematic from a pragmatist perspective. Until then, it might be sufficient to note that even this move might not be enough. What if Vulcan is actually transparent, made out of dark matter perhaps? What if, for this reason, we consider that interpretations of our cosmic model should not map Vulcan to any visible object? Then Newtonian mechanics would be safe. It seems that with enough imagination, and if a theory having at least one accurate model were enough, we could make any theory accurate, even considering the universe as a whole. In light of this, we might wonder why scientists did attempt to observe Neptune, instead of simply postulating that it is invisible.

This example from astronomy involves observable objects, in van Fraassen's sense of the term. However, the property of being observable does not play any particular role in this narrative (nor in our notion of model adequacy): observing a planet is just one way of accessing a property denoted by our model. The same difficulty can arise for any kind of theory or model. What is really needed to avoid this kind of difficulty is a distinction between ad-hoc and legitimate hypotheses. Assuming that Vulcan is transparent is ad-hoc, and therefore illegitimate (or at least problematic).

All these remarks are, of course, related to Duhem's observation that it is always possible to save a theory in the face of an empirical failure. Typical examples of Duhemian underdetermination involve the malfunctioning of apparatus. Arguably, verifying that apparatus are functioning properly amounts to verifying that our interpretation of the model in terms of the target (here, the interpretation of the data model) is correct. It is a matter of *relevance*. So, the right way of solving Duhemian problems is to put constraints on the relevance of models. This points to the idea that a model with Vulcan, or a model without Neptune are irrelevant from the point of view of Newtonian mechanics because Vulcan does not exist while Neptune does. This should be expressed by a general model of the solar system. In consequence, the second of our two options above should be favoured.

Accepting this, the move consisting in extending our models to incorporate more and more phenomena until we reach a model of the universe appears to be superfluous. We are not looking for a theory that merely has at least one adequate model, but for a theory that can produce a *relevant* and adequate model, and this can be assessed for local models. I will say more about what should be expected from relevance in order to avoid ad-hoc hypotheses shortly.

Another reason to reject the first option is that, as we have seen in the previous chapter (Sects. 3.3.4 and 3.4.1), the fact that a model is or is not licensed as a representation of its target by the epistemic community is responsive to empirical aims. The notion of communal licensing, which I borrow from Boesch (2017), has this implication in the case of scientific models (see Sect. 2.4.2). Licensing depends on factual knowledge or postulates concerning the target or its type. The notion of relevance is not limited to a correct interpretation of theoretical vocabulary: more is involved, in particular if we acknowledge that models enjoy a certain autonomy from theories. Domain-specific models are an important product of science: claiming that they merely play the instrumental role of showing that the theory has at least one adequate model for every situation or for the universe as a whole seems far-fetched. This means that relevance should not be as liberal as is implied by this first option.

4.3.3 *Unproblematic Failure as Irrelevance*

Let us now examine the second option in detail: the model without Neptune is irrelevant, which is why Newtonian mechanics should *not* be rejected. This option

assumes that a theory can evolve while remaining the same theory, because its conditions of application evolve, and that relevance (which interpretations of the theory are licensed in which contexts) can be sensitive to empirical inputs, in our case, the observation of Neptune and the lack of observation of Vulcan.

What could seem unclear is what kind of empirical inputs should be taken into account to assess relevance. As noted earlier, the risk of this option is failing to account for the possibility of misrepresentation. If all empirical data count for a model to be licensed, then relevance will collapse into adequacy. We could say, for example, that a model without Vulcan is *not relevant*, because it does not account for the trajectory of Mercury, and that a model with Vulcan is not relevant either because Vulcan was never observed, and in that case Newtonian mechanics would be safe (although it would be unable to represent the solar system!). This is not a desirable option. We would rather say that a model without Vulcan is perfectly relevant, because Vulcan cannot be observed, but inadequate, because it does not account for the trajectory of Mercury, and that this points to the inadequacy of Newtonian mechanics. Perhaps we could distinguish between two kinds of properties: the ones that matter for relevance (the number and initial positions of solid bodies?) and the ones that only matter for accuracy or adequacy (their full trajectories?). But as argued in Sect. 2.3.2, it is not clear that there is a principled way of doing so.

What I propose to solve this problem is to follow Lakatos (1978)'s insight that new theoretical posits introduced in order to save a theory should receive independent confirmation. What this means, in our framework, is that these posits could in principle be used in different experimental contexts, with the same or with different general models, and lead to empirical success. For instance, a general model of the solar system purports to be applicable to predict the position of any massive body that figures in the model at any time, and the contexts of application to which it applies should not be limited by ad-hoc restrictions to the ones where it proved successful. To say it differently, norms of relevance should be cross-contextual and unified. They should ensure a certain consistency in the application of the theory in various contexts. They are evaluated by means of one epistemic value in particular: *coherence*. This does not mean that we should be able to represent the universe as a whole, nor that models cannot incorporate domain-specific postulates. It only means that the way we represent various parts of the universe should be coherent overall. This idea is consistent with coherentist analyses of ad-hocness (Schindler 2018). Myrvold (2003)'s notion of unification could be used to formalise what is at stake. I will say more about this notion in Chap. 6 (Sect. 6.2.2). In this context, we could say that the model that should be licensed for particular applications is, among the coherent models, the one that performs best with regards to adequacy, or perhaps the one that achieves the best balance between coherence and adequacy. This model can fail to be adequate, and then, the theory is inadequate.

According to this picture, positing new entities such as Vulcan and Neptune amounts to putting into question the conditions of relevance of a theoretical model of the solar system, because this means considering that another incompatible general model is actually more relevant than the original one for the solar system. When

scientists are confronted with an experimental failure, they do not assume that their theory is inaccurate: they first consider that their theoretical model could be irrelevant. This is a way of “saving the theory”, but this hypothesis is awaiting confirmation. They then propose a new model that could be more relevant and more accurate at the same time: a model with Neptune or Vulcan, for instance. Since such posits are rightly understood as proposals for updating the way the theory is applied to the world in general, these posits should receive independent confirmation before the corresponding norm of relevance can be licensed. Scientists have to show the cross-contextual applicability of the norm to justify its legitimacy, and the way to do so is to use these norms in new representational contexts: for example, by attempting to observe Neptune using the new general model.¹¹ If this works, then the new norm is likely to be adopted, as in the case of Neptune. If it fails, as in the case of Vulcan, then maybe other norms of relevance will do (after Vulcan, scientists posited an asteroid belt to account for the trajectory of Mercury). Or maybe the first model was relevant after all, since no better model can be found, and the theory is inaccurate. But we do not know: we have an anomaly.

Coherence could a priori concern various levels, and so, various norms of relevance. Remember that norms of relevance specify the legitimate interpretations of models. They can roughly be understood as rules of interpretation: at the level of meta-norms, such as the fundamental laws of theory, a symbol should be interpreted in terms of a particular type of quantity (for example, acceleration) only if it respects structural constraints with other quantities (forces), and at the lower level, a model should be interpreted in terms of a particular type of target only if it has a particular structure (for example, respecting the Navier-Stokes equations for representing a fluid) (see Sect. 3.3.4). Yet other norms concern the legitimate operationalisations of theoretical terms (Sect. 3.4.2) and the combination of simple models into complex ones (Sect. 3.3.5). Coherence is the requirement that these rules of interpretation should be as unified as possible,¹² and that exceptions to the rules should be avoided as much as possible. In the case of compositional norms, coherence could imply that various complex models should use the same simple models when they apply to the same targets, and that these simple models should be licensed as well. This is not incompatible with a certain flexibility in the application of the theory to specific domains.

As we can see, this solution does not run into the same difficulties as the previous one. We do not need to consider more exhaustive contexts, up to the representation of the whole universe, to avoid trivialisation. All that is needed is that every context is addressed in a consistent way with other applications of the same model, and in this respect, the idea that Vulcan would be invisible is rightly

¹¹Note, again, that the fact that Neptune is observable in van Fraassen’s sense plays no particular role in this illustration. The idea is to access a posited entity by various empirical means: here, the direct observation of Neptune by means of a telescope and the indirect observation of its effect on the trajectory of Uranus.

¹²An example of this is the way models of protein integrate various experimental models, corresponding to different experimental contexts, mentioned in footnote 4.

understood as problematic, because it has very limited cross-contextual coherence, since the corresponding norm only applies to one specific object.

Remember that a model of the solar system is a particular case in that a specific object is represented by the general model, but in general, a theoretical model, such as the model of a pendulum, will be applicable to various instances of a particular type, and the specificities of an instance are not extensible to other contexts of use. This means that relevance does not collapse into adequacy, because some empirical inputs are specific to particular experimental contexts. They cannot be involved in cross-contextual norms, and they cannot appear in a general model. Yet they matter for accuracy. We could expect that in general, the aspects that will receive cross-contextual confirmation will concern the existence of objects of a certain type in specific targets or the categorisation of types of phenomena (how to model them, which value model parameters should take), whereas the context-specific aspects will concern dynamical quantities or contingent states. The former aspects would matter for relevance, because by their nature, they should affect other contexts of use of the theory. This gives us a sense of what a *type of situation*, compatible with various possible states, could be, which will be useful in the next chapter.

If we accept this stricter notion of relevance, we can say that a theory is adequate if all its relevant models are: we do not need to say “at least one model”, as in the first option. The idea that all relevant models should be adequate for the theory to be adequate is *prima facie* consistent with scientific practice: an experimental failure calls for a revision of our assumptions in one way or another, and it generally affects which models are considered relevant for representing particular types of targets by the community. Having two models making incompatible predictions should be considered problematic: one has to go, which would not be implied if we considered that a theory only needs to have one adequate model to be adequate. Furthermore, scientists sometimes do not create a model to represent a situation, but rather create a situation that the model can represent (see Sect. 2.3.3). If the experiment fails, this is a problem for the theory. This shows that scientists want all, and not only some, models of a theory to be empirically adequate whenever they are relevant.

Accordingly, we are reaching our final definition of theoretical adequacy:

Definition 4.4

$$adequacy(T) := (\forall M \in T)adequacy(M)$$

Unpacking the modal version of the adequacy function that applies to models gives us the definition that characterises modal empiricism:

Definition 4.5

$$adequacy_m(T) := (\forall M \in T)(\forall S)\Box_S(\forall C \in affordance(S))(\forall m \in relevance(M, C))accuracy(m)$$

This, completed with a notion of coherence for the relevance function, is the main criteria by which theories are accepted and rejected according to modal empiricism, and scientists have good reasons to assume that currently accepted theories are adequate in this sense.

4.3.4 *Theory Evolution and Theory Change*

Our conclusion in this section is that the best way of accounting for the impact of experimental failures on scientific theories is to consider that the norms of relevance guiding which models are or are not apt to represent which targets are rather strict, but revisable, and that they have to do with cross-contextual coherence.

The resulting picture of the functioning of science is quite compatible with Kuhn (1962)'s account (except for his commitment to incommensurability). More specifically, the idea that norms of relevance are revisable allows us to consider two ways of updating our representation of the world in the face of an experimental failure: either make the theory evolve by updating norms of relevance, that is, by constructing new theoretical models conveying these norms, or change the theory by switching to a completely different family of general models. These two ways follow Kuhn's distinction between normal science and scientific revolutions, or Lakatos's distinction between the core of a theory and its protective belt.¹³

The first way of updating our representations corresponds to what Kuhn calls normal science. Puzzles are solved; new theoretical models are constructed to account for new phenomena; the theory is extended to new domains. Before first observing Neptune, the trajectory of Uranus constituted a puzzle. Scientists could fear that their new model was not relevant, but it turned out to be: a problem was successfully solved.

The second way of updating our representation of the world corresponds to what Kuhn calls a scientific revolution, which can happen when the accumulation of anomalies calls for a more drastic revision. In our framework, this should occur when there is no way of achieving accuracy in various circumstances while maintaining the cross-contextual coherence of norms of relevance.

This means that the unificatory power of theories is an important component of empirical adequacy, and it is, for a modal empiricism, part of the aim of science. Incidentally, I would say that the radical pluralism put forth by some philosophers of science should be tempered (see Sect. 2.3.1). After all, fundamental physics seems to progress towards greater unification, from Newton's theory unifying celestial and terrestrial mechanics to the standard model of particles unifying most fundamental forces. To be sure, physics has also developed towards more specialised disciplines, and pluralists are right to emphasise that domain-specific models have a certain

¹³One could have a finer picture where various levels in a hierarchy of norms can be revised, the theory being the highest level, as suggested in Sect. 3.4.1.

autonomy from fundamental theories. But there is no contradiction here: coherent meta-norms constraining the interpreted structure of theoretical models can coexist with domain-specific applications of these constraints.

One could question the idea that science *should* aim at unification. The idea makes sense for a realist, but what is the added value of unification for an empiricist? Why entertain such an aim if domain-specific models can be successful without having to be coherent with the models of other domains? An answer to this question can be found in the work of philosophers who associate a certain form of understanding with unification. Friedman (1974) has proposed that scientific explanations provide understanding because they reduce the number of brute facts that we have to accept in order to account for a variety of phenomena. Kitcher (1989) has improved on this theory, and proposed that unification works by allowing us to apply the same *patterns of derivation* in various circumstances. This notion of pattern of derivation is interesting, since it preserves a relative autonomy for domain-specific explanations, and it is not far from the notion of coherent norm of relevance. Remember that norms of relevance are functions from context to interpretation, so that the general models conveying them are filled in by contextual inputs in the same way as a pattern. So, unification has the virtue of cognitive economy, and perhaps this gives us the ability to withstand complex situations on the basis of our knowledge of simpler ones, which could be a useful strategy for extending a theory to new domains. After all, our possible experiences are not always partitioned into hermetic domains.

Another question is: why think that this unification is achievable? This question will be addressed in the following chapters.

In this picture, the aim of science is to produce theories that are both accurate and coherent in the way they are applied. Such theories unify various possible observations in a consistent conceptual scheme, and this is what the definition of empirical adequacy proposed above expresses. This definition thus incorporates two cognitive values: conformity with experience and coherence (which I take to imply unificatory power). However, this notion of empirical adequacy might not fully capture the aim of science. There are trivial ways of being empirically adequate for a theory, according to this definition. One way is to have no model at all, or models that are never relevant. Another way is to have models that are too permissive, and that bring no constraints on what is or is not possible. This means that other desiderata than empirical adequacy must be considered in order to properly account for the aim of science.

One such desideratum is *scope*. Ideally, the range of situations for which the theory has a relevant model should be maximal. Perhaps we could say that an ideal theory should be universal, that is, that it could represent any situation in the world, for any context (although arguably, this concerns physics more than the special sciences). Such a value can be formalised in the present framework as follows:

Definition 4.6

$$\begin{aligned} & \text{universality}(T) := \\ & (\forall S) \Box_S (\forall C \in \text{affordance}(S)) (\exists M \in T) (\exists m \in \text{relevance}(M, C)) \end{aligned}$$

This aim is very idealistic, but going in its direction could be desirable. Another desideratum, which might be more attainable, is *informativeness*. The theory should be able to tell us exactly what is or is not possible in a particular situation. We could say that an ideal theory would be perfectly informative if its models would not permit any impossible state of affairs. Formalising this notion requires unpacking the *accuracy* function in terms of the actual state of the situation belonging to the states permitted by an interpreted model. For any conceivable context in any possible situation, and for any state permitted by any licensed interpretation of a theoretical model in this context, there should be a possible way for the situation to be that corresponds to this permitted state. This would give the following definition (where $\text{state}(C, S)$ denotes the state of situation S from the perspective of a context C):

Definition 4.7

$$\begin{aligned} & \text{informativeness}(T) := \\ & (\forall M \in T) (\forall S) \Box_S (\forall C \in \text{affordance}(S)) (\forall m \in \text{relevance}(M, C)) \\ & \quad (\forall s \in \text{permitted}(m)) \Diamond_S \text{state}(C, S) = s \end{aligned}$$

Even if our definition of empirical adequacy is not a complete characterisation of the aim of science, it still captures its main component, since scope and informativeness would not make much sense without empirical adequacy. We will see in Chaps. 6 and 7 that it can also help us respond to the main arguments in the debate on scientific realism.

4.4 Comparison with van Fraassen's Account

In the previous section, we arrived at the definition of empirical adequacy that characterises modal empiricism. It roughly states that a theory is empirically adequate if all its models correctly capture the way natural constraints on possible phenomena express themselves in the contexts where these models can be interpreted, that is, the way these natural constraints limit possible manipulations and observations of actual situations.

There are three main differences between the proposal for empirical adequacy that we have reached in the previous section and van Fraassen's original definition:

- this new definition of empirical adequacy is situation-based, while van Fraassen's is universe-based: it quantifies over all models, situations and contexts, while van Fraassen's involves one model of the theory that could represent the whole universe;
- this definition is based on the pragmatic, community-centred notion of relevance, while van Fraassen's is based on the objective notion of observable;
- this definition is modal, while van Fraassen's is not.

I will argue that these three differences make this new definition more appropriate to serve as an axiological notion characterising the aim of science. I will also show that these three aspects are related: they constitute a coherent "package", which is distinctively pragmatist. Let us examine these aspects one by one.

4.4.1 *Situations Versus the Universe*

Our situation-based account of empirical adequacy considers the ability of a theory to represent accurately any situation in the universe, by means of various models. In contrast, van Fraassen's universe-based definition considers the ability of a theory to represent accurately the whole universe by means of a single model. Perhaps van Fraassen would accept this modification, given his recent analysis of the contextual aspects of representation. In any case, a situation-based account is more modest and connected with scientific practice.

Focusing on a putative model of the universe for philosophical analysis is a common practice in philosophy of science, in particular in metaphysics (see for example Belot 2016; Greaves and Wallace 2014). It is somehow related to what I have called a theory-first approach in the introduction of Chap. 3. A universe-based approach can be seen as a way of ensuring a certain axiological unity for science, which is apparently threatened by an understanding of theories as collections of models applied contextually. The disconnect between this approach and scientific practice can be seen in the fact that in general, scientists do not act as if they were willing to integrate all their experimental results into an all-encompassing model. For example, they do not necessarily write down the exact position and time of the phenomena they record, nor do they reproduce past experiments over and over again, so as to ensure that their models still make correct predictions in new places and times. The fact that a given experiment is successful seems sufficient to confirm a theory or hypothesis, irrespective of its place and time. At most, scientists could require that independent teams reproduce a particular experiment, so as to eliminate implicit biases in the way it was carried out.

This idea of providing a model of the universe seems more applicable to the theories of physics than to other disciplines. But even in physics, no one has ever seen a model that would contain all observable phenomena in the universe, from the tiniest dust particle to the biggest structure. Cosmological models, with their assumption that the universe is homogeneous, are a far cry from it. If the universe

is infinite and irregular, such a model is not even constructible by finite means. Unsurprisingly, the examples that serve as illustrations of empirical adequacy in the literature, including in van Fraassen's writing, are always examples of theories as applied to particular situations through particular models, not to the universe as a whole. So why invoke such an idealisation?

The aforementioned observations are a bit unfair, though. Admittedly, van Fraassen does not claim that the aim of science is to produce a model of the universe. He only claims that it is to produce a theory that would have such a model. So, the fact that scientists do not feel the urge to record the exact place and time of their experiments might not be relevant after all. Perhaps any empirical success could count as an indirect confirmation that our theories might have a model of the universe (notwithstanding the fact that this model might not be constructible).

However, this idea rests on a set of implicit postulates that do not all go without saying. For one, it seems to assume an eternalist metaphysics, according to which there is a determinate set of past, present and future observable phenomena. Without this assumption, no theory could be *presently* empirically adequate, or only about past phenomena, which would contradict van Fraassen's definition. Some practising scientists would be sympathetic to this eternalist idea, but empirical adequacy is not concerned with the aim of particular scientists, and there seems to be no *a priori* reason to count this metaphysical assumption as a prerequisite to make sense of scientific practice.

The idea that local successes confirm the ability of the theory to produce a model of the universe also assumes a sort of compositionality principle. In order to infer, from the accuracy of a particular model of the theory in a given context, that the same theory could produce an adequate model of the universe as a whole, one has to assume that this particular model is contained in the model of the universe, or at least, that its empirical sub-structures are contained in the empirical sub-structures of the model of the universe. This should hold for any particular model of any situation. In other words, one has to assume that any part of a model of the universe provided by the theory is also a model of the same theory, or can be turned into one while preserving the empirical sub-structures. This would explain why the exact place and time of an experiment do not matter.

Arguably, this compositionality principle is more or less respected by contemporary theories to some extent. Physicists, for example, have tools at their disposal to combine models into larger ones, or to split models into components, and these combinations and components are all models of the same theory. This principle is similar to what I have called compositional norms. However, if we accept this principle, the advantage of focusing on a model of the universe is not clear, because having an accurate model of the universe implies having accurate models for any situation in the universe.

In this respect, a situation-based account of empirical adequacy, which focuses directly on the constraints that the theory as a whole brings to particular models in particular contexts instead of referring to a putative cosmic model, brings more connections to scientific practice, and also a form of modesty. Indeed, a focus on

particular models and situations involves an inversion of quantifiers.¹⁴ Simplifying a bit, a universe-based account assumes that there is one model such that all situations “fit inside” ($\exists M \forall S$) whereas a situation-based account assumes that all situations are such that they fit inside a model ($\forall S \exists M$, neglecting the difference between “at least one model” and “all relevant models”). The second is logically weaker, in particular if not all arbitrary conjunctions of situations count as potential targets of representation. Maybe some of these arbitrary conjunctions are not potential targets of representation, even assuming that there are norms of composition. This could be the case with objects too far apart in an expanding universe to ever have a common causal future. This could also be the case if the way a theory is applied to various types of phenomena is not as unified as one could expect (Wilson 2013). The whole universe, in all its tiny details, might not be a potential object of representation, if only because the users of representation are part of it, whereas one would expect them to be external to the represented situation.

Particular, bounded situations are our primary access to reality, and parsimony enjoins us to refrain from postulating that our theories could produce accurate models of things that we could not necessarily represent or experiment on, such as the whole universe. This is a way of staying closer to actual scientific practice. This also gives all models of the theory a potential role, and not only one of them, which does justice to the spirit of the semantic conception of theories. Our definition of empirical adequacy in terms of situations is therefore more suitable for an empiricist willing to emphasise experience as the primary source of justification for knowledge.

4.4.2 *Relevance Norms Versus Observables*

As explained in Sect. 4.3.2, one possible motivation for a universe-based account of empirical adequacy could be to impose a constraint of coherence on the way the theory “saves the observable phenomena”. In the account presented in this chapter, this role is played by norms of relevance and their cross-contextual unity. We could say that modal empiricism trades an extensional notion of coherence and unity for an intensional one: the unity of the theory is provided by its ability to be accurate in all possible contexts.

Another (but related) role played by norms of relevance is to ensure some kind of objectivity in the way a theory is connected to experience. In the case of science, this connection is often mediated by models of experiments, and as suggested in the previous chapter (Sect. 3.4.2), these models presumably have their own norms of relevance, associated with licensed operationalisations of the theoretical vocabulary. These norms have to do with cross-contextual coherence as much as theoretical norms do. They concern the stability and reproducibility of experimental results

¹⁴I am indebted to an anonymous referee for this observation.

and the unification of various operationalisations when they are equivalent, as well as the coherence with theoretical norms, in particular when theories are used to explain the functioning of measuring apparatus.

Using the distinction between appearances and phenomena employed by van Fraassen, we could say that experimental norms have to do with the way experimenters manage *appearances* and connect them with *phenomena*, and they are distinct from the norms connecting phenomena to theoretical models. However, this framing in terms of norms is deflationary in spirit: I do not wish to assume that phenomena are real, observable entities. They merely correspond to the cross-contextually stable categories in use in scientific experimentation. This constitutes another important difference between modal empiricism and constructive empiricism: van Fraassen assumes that real, observable phenomena exist, and they play a crucial role in his constructive empiricism. In his account, the connection between theory and appearances is not mediated by pragmatic, communal norms of experimentation, but by observable phenomena. It is important to note that according to van Fraassen, observable phenomena are distinct from appearances. Appearances correspond to how phenomena look from a given perspective. For example, “Mercury’s motion is an observable phenomenon, but *Mercury’s retrograde motion* is an appearance” (van Fraassen 2008, p. 287, emphasis in the original).¹⁵

Relying on pragmatic, communal norms of relevance instead of an objective notion of observable phenomena can be seen as a virtue given the huge amount of criticism that the notion of observable has received (Psillos 1999, ch. 9) (Churchland and Hooker 1985, ch. 2) (Alspector-Kelly 2004; Buekens and Muller 2012; Gava 2016). It can be argued, for instance, that the phenomena that are accounted for by scientific models are rarely observable in van Fraassen’s sense, and that whether they are or not does not seem to matter for scientific practice, so being a realist about phenomena (and not only about appearances) is in tension with empiricism. Another crucial advantage of the notion of relevance is that it can accommodate a sensitivity to contexts and purposes in a way that the notion of observable cannot. In particular, with the notion of observable, contextual purposes must play a merely selective role in representation, or, at most, a productive role *prior* to representation,

¹⁵Note that our account of representation implies several layers. A model of experiment connects direct observations or “appearances”, which constitute the first layer, with their theoretical descriptions by interpreted models, the second layer, but these theoretical descriptions are still apprehended from a particular context associated with a focus on discrete properties of interest. A theoretical model connects these contextual descriptions to cross-contextual, indexical representations that could apply in various contexts, which constitutes the third layer. A fourth layer is constituted by theoretical laws, which unify various general models. The first three layers presented here more or less match van Fraassen’s notions of appearances, observable phenomena and unobservable phenomena respectively. However, this match is not perfect, because the way these layers are distinguished is markedly different. In particular, the notion of being observable plays no role beyond appearances in the account used in this work, which is the point discussed here.

that of “enlarging the observable world”, while with relevance, they could play a constitutive role in the representation relation (see Sect. 2.3.3).

Even if it dispenses with the notion of observable, modal empiricism is still firmly empiricist. According to modal empiricism, the ultimate benchmark for empirical adequacy is the accuracy of particular models in particular experimental situations, and this accuracy is evaluated with regard to empirically accessible characteristics of these situations. Relevance merely has to do with scientists' management and interpretation of “appearances”, in van Fraassen's terminology, unification and experimental stability being the main criteria. Modal empiricism crucially relies on appearances: it merely bypasses the problematic intermediate level constituted by observable phenomena in van Fraassen's account.

One worry could be that metaphysical posits are smuggled in by scientists and their norms of relevance, whereas they would be blocked by a reliance on observable phenomena. Models could be licensed for metaphysical reasons: because they posit the right kind of unobservable entities for example, where these would be determined by an inference to the best explanation. Since norms of relevance determine conditions of accuracy, this would make empirical adequacy equivalent to theoretical truth, because a theory could only be empirically adequate if it correctly describes unobservable entities. However, this is only true assuming, as the realist generally does, that inference to the best explanation is truth-conducive.

Metaphysical considerations, or metaphysical prejudices inherited from our ancestors, may sometimes play a heuristic or psychological role in science. I do not deny that they could affect the relevance norms that are actually used by scientists for interpreting theoretical models. However, the important point is that they do not play a normative role, in the sense that complying with them would be part of the aim of science. New theories often contradict our metaphysical prejudices. Our pre-theoretical intuitions about time are contradicted by relativity theory, when interpreted realistically, for example. This is no reason to reject these theories. According to modal empiricism, if a theory can “save the appearances” in a more unified way than another, then it should be preferred, whatever metaphysical intuitions were used to construct the theory. The realist could argue that an ability to save the appearances in a unified way is an indicator of truth, so that science, by aiming at empirical adequacy, actually achieves truth. I doubt it, but this question is different from the axiological question that occupies us here (this issue will be addressed in Chap. 6).

4.4.3 *Modality Versus Extensionality*

The last difference between my definition and van Fraassen's is the modal aspect. This might be the most controversial difference for an empiricist. It is also related to the previous issue, in that a reliance on observable phenomena can be seen as a means of extending empirical adequacy to situations that are not experimented upon without invoking modalities. But as we have seen in Sect. 4.2.4, modalities are

required if experimentation is, in general, active: perhaps *observable* phenomena are actual even if never observed, but it is much less clear that *producible* phenomena would be actual even if never produced.

Endorsing modalities also makes the whole position more coherent for different reasons.

First, this idea allows us to avoid problems with the notion of observable, which is apparently modal. Van Fraassen claims that “observable” is an objective property of actual phenomena, about which our theories can inform us (although this notion is theory-independent, so as to avoid a vicious circularity (van Fraassen 1980, p. 57)). But such an objective notion of observable does not appear explicitly in any scientific theory. It is not obvious that the notion can be explicated without assuming that the theories that inform us about what is observable have modal import, in particular if we want these theories to tell us whether types of phenomena that were never observed so far are observable (Ladyman and Ross 2007, ch. 2.3.2.3) (on this controversy, see Rosen 1994; Ladyman 2000; Monton and van Fraassen 2003; Ladyman 2004; Dicken and Lipton 2006).

The idea that empirical adequacy is about saving possible observations can also help us make better sense of scientific practice, following van Fraassen’s own line of reasoning. Van Fraassen extends empirical adequacy to all observable phenomena to account for the motivation of scientists when they experiment on new phenomena (Monton and van Fraassen 2003). His argument can be reconstructed as follows: if only actually observed phenomena mattered for empirical adequacy, scientists could just as well observe nothing, and they would have achieved their aim. This is not what they do. Scientific practice only makes sense if scientists want to save all observable phenomena, and take actually observed phenomena to be representative of merely observable ones.

This argument can be readily extended to merely possible phenomena. If only *actual* phenomena mattered, scientists could just as well refrain from creating new phenomena that would probably not occur without their interventions. But scientists rarely draw a distinction between natural and artificial phenomena, the latter being those that would not happen if it were not for their interventions. They do not perform a passive observation of all the heterogeneous situations they stumble upon, which is what they could do if they wanted their theories to be empirically adequate for all actual situations. Instead, they actively create situations of interest, as if they were exploring possibilities. Physicists sometimes carry out thought experiments, such as the EPR-type experiment involved in the derivation of Bell’s inequalities, while following purely theoretical considerations. What they have in mind, it seems, are intriguing consequences or principles of their theories, and they are willing to create experimental situations so as to explore these possible consequences, regardless of whether they occur naturally. They often control target systems and vary parameters as if they were exploring various possibilities. All this is strongly related to the remarks in Sect. 4.2.4 about the fact that experimentation generally involves finely controlled interventions and exotic configurations. Testing theories against various possibilities seems to be part of the “rules of the game” of scientific practice (see Ladyman and Ross 2007, p. 110 for a similar argument).

Van Fraassen could claim that by doing so, scientists are merely trying to improve their theories for further application to natural phenomena (for example, by fixing the value of missing constants). However, in the case of, say, the implementation of EPR-type experiments, this seems far-fetched. No theoretical constant is involved in this case, and arguably, if no violation of Bell's inequalities had been observed by Aspect et al. (1982), the right conclusion would have been that quantum mechanics is not empirically adequate, or at least, that its foundations and interpretation (its norms of relevance) must be adapted in consequence. When this type of experiment fails, the natural conclusion is that the theory would have been inadequate even if the experiment had never been performed. Scientists did not make the theory inadequate by their intervention; they merely realised that it was, so empirical adequacy must concern merely possible situations. But even rejecting this natural interpretation, the fact that scientists took the risk of implementing this experiment shows that they wanted to make sure that their theory would correctly account for this possibility.

One could argue that the notion of possibility involved is merely epistemic: for all we know, such situations might actually occur somewhere in the universe, so we should make sure that our theories account for them. The problem with this response is that the likelihood of this occurrence does not seem to play any role in the motivation of scientists.

In light of this, there seem to be good reasons to think of empirical adequacy, taken to be a criterion for ideal empirical success, as a modal notion: it concerns the adequacy of the theory to represent or predict what would be observable in some specific experiments that might not be actual, given the natural constraints on what is possible for us to do and observe in this world. The aim of science is to produce theories that would be applicable with success in all possible circumstances, without presuming anything about what is actual. Actual experimental results are taken to be representative or indicative not only of unobserved actual phenomena, but also of mere possibilities. Empirical adequacy is an ampliative notion. In this respect, the fact that these unrealised possibilities are not actually realised does not matter any more than the fact that unobserved phenomena are not actually observed: the ampliative move is of the same nature (this line of reasoning will be developed in the next chapter).

There could be objections to this account regarding the commitment to the existence of worldly possibilities on which it rests. One issue has to do with the principled impossibility of modal knowledge. I will alleviate this worry in the next chapter. Another objection could be that the resulting position is less metaphysically parsimonious, because of this commitment to natural modalities. The aim that modal empiricism attributes to science is certainly more ambitious than the one attributed by constructive empiricism, because it extends to the realm of possibilities, but when it comes to metaphysical commitments, I am not sure that modal empiricism really is less parsimonious.

As already noted, constructive empiricism seems to assume an eternalist metaphysics: our theories must ideally match a distribution of past, present and future observable phenomena, and these phenomena must all exist, or our theories could never be empirically adequate *now*. Arguably, taking position in the debate on the

metaphysics of time in order to account for the rationality of science is problematic. The definition of empirical adequacy that I have suggested quantifies over actual situations as well, but it can remain agnostic on which situations (past, present or future) are actual. Our theories could still correctly account for future possibilities, understood as possible ways for present situations to evolve.¹⁶ So, in the end, rather than opting for a less parsimonious position, all we have done is trade an implicit eternalist metaphysics for natural modalities.

4.5 Modal Empiricism as Pragmatism

The aim of this chapter was to examine how one shall understand empirical adequacy in the context of the account of scientific representation provided in the previous chapter. The strategy adopted followed a bottom-up approach: empirical adequacy is achieved as a result of two ampliative moves, first from model accuracy in context to model adequacy in general, then from model adequacy to theory adequacy.

This analysis resulted in several important conclusions. One of them is that empirical adequacy must be a modal notion, in order to both extend to non actual contexts of use and account for the interventionist aspects of experimentation. Accepting that empirical adequacy is a modal notion is the only way of making sense of scientific practice as a rational activity: the fact that scientists implement exotic experimental configurations to test their theories would be unintelligible without it. The other important conclusion we reached is that the relevance of a model for representing phenomena of a particular type is mainly a matter of cross-contextual coherence in the way the model is applied, and that norms of relevance are substantial, revisable and sensitive to empirical inputs. This cross-contextual coherence is strongly related to the unificatory power of models and theories.

The resulting notion of empirical adequacy, which characterises the empiricist position defended in this book, is the following one: a theory is empirically adequate if all its relevant applications, including in counterfactual contexts, would be successful. In other words, an ideal theory should correctly account for all possible manipulations and observations we could make within a domain of application, assuming these possibilities are limited by natural constraints. One could say that an empirically adequate theory gives us a reliable map of what is or is not possible in this world, for a given domain of activities. According to modal empiricism, the aim of science is to produce theories that are empirically adequate in this sense.

This notion of empirical adequacy differs from van Fraassen's in three respects. It is situation-based, in that it does not involve any reference to the whole universe. It

¹⁶In a sense (putting aside issues that have to do with the notion of observable), modal empiricism is not *much* more than constructive empiricism made compatible with the openness of the future, which happens to be more in tune with scientific practice.

is pragmatic, because it relies on norms of representation rather than on a notion of observable phenomena. And as we have just seen, it is modal. One could say that we have on one side a definition that is centred on an ideally complete representation of a cosmic state of affairs, and on the other, one that is centred on an ideally successful conceptual scheme, constituted of norms of representation that could allow us to withstand any possible situation. It should be clear, for this reason, that the definition of empirical adequacy proposed in this chapter does not bring empiricism closer to scientific realism, despite its commitment to natural modalities. It is more realist when it comes to modalities, but has more instrumentalist flavours in other respects. What it does is bring empiricism closer to pragmatism.¹⁷

We now have a good understanding of what modal empiricism is. What remains to be done is to show that modal empiricism fares better than other positions in the debates on scientific realism. An important step in this direction consists in showing that modal empirical adequacy is an achievable goal, and in particular, that the natural modalities to which the modal empiricist is committed do not correspond to an “inflationary metaphysics” that would make the whole position subject to a problem of underdetermination by experience. Arguably, the alleged impossibility of modal knowledge is the main reason for empiricists to reject the existence of natural possibilities. In the next chapter, I will argue that this reason is not acceptable, because modal knowledge is perfectly achievable.

References

- Alspector-Kelly, M. (2004). Seeing the unobservable: Van Fraassen and the limits of experience. *Synthese*, 140(3), 331–353. <https://doi.org/10.1023/B:SYNT.0000031323.19904.45>
- Aspect, A., Dalibard, J., & Roger, G. (1982). Experimental test of Bell's inequalities using time-varying analyzers. *Physical Review Letters* 49(25), 1804–1807. <https://doi.org/10.1103/PhysRevLett.49.1804>
- Belot, G. (2016). Undermined. *Australasian Journal of Philosophy*, 94(4), 781–791.
- Boesch, B. (2017). There is a special problem of scientific representation. *Philosophy of Science*, 84(5), 970–981.
- Buekens, F. A. I., & Muller, F. A. (2012). Intentionality versus constructive empiricism. *Erkenntnis*, 76(1), 91–100. <https://doi.org/10.1007/s10670-011-9348-1>
- Bueno, O. (1997). Empirical adequacy: A partial structures approach. *Studies in History and Philosophy of Science Part A*, 28(4), 585–610.
- Churchland, P., & Hooker, C. (1985). *Images of science: Essays on realism and empiricism*. Chicago: University of Chicago Press.
- Dicken, P., & Lipton, P. (2006). What can Bas believe? Musgrave and Van Fraassen on observability. *Analysis*, 66(291), 226–233.
- Friedman, M. (1974). Explanation and scientific understanding. *The Journal of Philosophy*, 71(1), 5. <https://doi.org/10.2307/2024924>

¹⁷Note, in this regard, that Peirce, one of the main figure of American pragmatism, endorsed natural modalities. The connection between pragmatism and modal empiricism will be elaborated in the concluding chapter of this book.

- Gava, A. (2016). Why van Fraassen should amend his position on instrument-mediated detections. *Analysis and Metaphysics*, 15, 55–76.
- Greaves, H., & Wallace, D. (2014). Empirical consequences of symmetries. *The British Journal for the Philosophy of Science*, 65(1), 59–89. <https://doi.org/10.1093/bjps/axt005>
- Kitcher, P. (1989). Explanatory unification and the causal structure of the world. In P. Kitcher & W. Salmon (Eds.), *Scientific explanation* (Vol. 8, pp. 410–505). Minneapolis: University of Minnesota Press.
- Kratzer, A. (2019). Situations in natural language semantics. In E. Zalta (Ed.), *Stanford encyclopedia of philosophy, metaphysics research lab*. Stanford: Stanford University.
- Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Ladyman, J. (2000). What's really wrong with constructive empiricism? Van Fraassen and the metaphysics of modality. *British Journal for the Philosophy of Science*, 51(4), 837–856.
- Ladyman, J. (2004). Constructive empiricism and modal metaphysics: A reply to Monton and Van Fraassen. *British Journal for the Philosophy of Science*, 55(4), 755–765.
- Ladyman, J., & Ross, D. (2007). *Every thing must go: Metaphysics naturalized*. Oxford: Oxford University Press.
- Lakatos, I. (1978). *The methodology of scientific research programmes*. Cambridge: Cambridge University Press.
- Mitchell, S. D., & Gronenborn, A. M. (2015). After fifty years, why are protein X-ray crystallographers still in business? *The British Journal for the Philosophy of Science*, axv051. <https://doi.org/10.1093/bjps/axv051>
- Monton, B., & van Fraassen, B. (2003). Constructive empiricism and modal nominalism. *British Journal for the Philosophy of Science*, 54(3), 405–422.
- Myrvold, W. C. (2003). A Bayesian account of the virtue of unification. *Philosophy of Science*, 70(2), 399–423. <https://doi.org/10.1086/375475>
- Psillos, S. (1999). *Scientific realism: How science tracks truth. Philosophical issues in science*. London, New York: Routledge.
- Rosen, G. (1994). What is constructive empiricism? *Philosophical Studies*, 74(2), 143–178.
- Ruyant, Q. (2017). *L'empirisme Modal*. PhD Thesis, Université Catholique de Louvain.
- Schindler, S. (2018). A coherentist conception of ad hoc hypotheses. *Studies in History and Philosophy of Science Part A*, 67, 54–64. <https://doi.org/10.1016/j.shpsa.2017.11.011>
- Suárez, M. (2005). The semantic view, empirical adequacy, and application (concepción semántica, adecuación empírica y aplicación). *Crítica*, 37(109), 29–63.
- van Fraassen, B. (1980). *The scientific image*. Oxford: Oxford University Press.
- van Fraassen, B. (1989). *Laws and symmetry* (Vol. 102). Oxford: Oxford University Press.
- van Fraassen, B. (2008). *Scientific representation: Paradoxes of perspective* (Vol. 70). Oxford: Oxford University Press.
- Wilson, M. (2013). What is “Classical Mechanics” anyway? In R. Batterman (Ed.), *The Oxford handbook of philosophy of physics* (p. 43). Oxford: Oxford University Press.

Chapter 5

Situated Possibilities, Induction and Necessity



Abstract Modal empiricism is the position according to which the aim of science is to produce theories that correctly account for all possible manipulations and observations we could make in their domain of application in a unified way. This chapter analyses the kind of modalities to which modal empiricism is committed, and responds to sceptical arguments against the possibility of modal knowledge. I defend an inductive epistemology for situated modalities, and I show that no practical underdetermination could affect general modal statements without thereby affecting the corresponding universal generalisations. This account is shown to fit well with scientific practice, thus vindicating modal empiricism as an account of the aim of science. The debate on what distinguishes laws of nature from accidental generalisations is briefly addressed. I suggest that law-like statements are just those that are justified by induction, and that they have a modal force. However, they are weaker and more diverse than the laws of nature as generally construed.

5.1 The Sceptical Challenge

We concluded the previous chapter with a characterisation of modal empiricism as the position according to which the aim of science is to produce theories that would be successful in all their applications, including counterfactual ones. This means, roughly speaking, theories that can account for all possible observations and manipulations we could make within a domain of application. It is now time to take a more argumentative approach, and to defend modal empiricism against its various opponents in the debate on scientific realism, in particular with regard to the question of justification.

The first type of opponent, which will be the focus of this chapter, is constituted of traditional versions of empiricism that would reject the modal component of modal empiricism. This modal component is expressed by the reference to all *possible* applications of a theory. This specificity of modal empiricism makes the position more ambitious than traditional versions of empiricism, and one reason to prefer a more traditional empiricism could stem from scepticism towards modal knowledge.

We seem to know a great deal about what is or is not possible in this world. For example, I believe that I know that glass is fragile, and that if I had let a glass vase fall from my window 1 min ago, it would have broken. I also know, from my physics classes, that it would have accelerated at a rate of about 9.8 m/s^2 : in normal conditions such as these, an object in free fall could not accelerate at a much lower or higher rate. All this, it seems, is common knowledge that came to be known by experience (mine or others'), not by merely reflecting on the meaning of words. Thus, modal knowledge is very mundane. But it poses an epistemological challenge: how could we know, from our experience, anything about unrealised possibilities?

One answer to this question, the one typically adopted by empiricists, is that we cannot. Hume notoriously observed that no necessary connections are found in experience. Other authors have argued that possible worlds, if there are such things, are causally disconnected from the actual world: we have no "modal telescope" to observe unrealised possibilities. Their conclusion is that we have no modal knowledge: at most, we can capture regular patterns in past phenomena, and anticipate future phenomena on this basis. The laws of nature, dispositions or causal relations either supervene on the mosaic of actual facts, or they only exist in our minds.

If these sceptical arguments are correct, then the notion of empirical adequacy presented in Chap. 4 becomes dubious. According to this notion, the "truth-makers" of empirical adequacy are possibilities in the world, including unrealised ones, and these possibilities are not mere conceivable states of affairs, but alternative ways actual situations of the world could be in virtue of natural constraints on possibilities. How could we know if a theory is empirically adequate in this sense without any epistemic access to unrealised possibilities? It does not make much sense to assign an aim to science if this aim is not achievable, and not even approachable as an ideal. One cannot play impossible games. This could be a reason to endorse more traditional versions of empiricism that do not incorporate this modal component, such as constructive empiricism.

In this chapter, I wish to defend modal empiricism against objections of this kind by proposing an empiricist-friendly epistemology for modalities. I will argue that it is rationally justified to assume that a theoretical model is empirically adequate in a modal sense on the basis of empirical evidence. I am indeed convinced that modal knowledge in general, and modal empirical adequacy in particular, are perfectly achievable, assuming that one empiricist-friendly mode of inference is warranted, namely induction. This is true at least considering one type of possibilities, the one involved in our definition of empirical adequacy: situated possibilities.

Induction is an ampliative mode of inference by which one infers, from a sample of exemplars of a given type displaying a regularity, that this regularity holds for a larger set: the set of all objects of this type. One received view has it that induction on actual experiences is insufficient to justify relations of necessity. It can only justify universal generalisations. As a consequence, empiricist-minded philosophers, even when they accept induction, tend to be sceptical about natural modalities, which they consider too metaphysical, while optimists about modalities tend to accept in their epistemology ampliative modes of inference that go beyond induction, typically,

inference to the best explanation. In this chapter (which is based on Ruyant (2019)), I challenge this received view by arguing that relations of necessity can be justified by induction, so long as we accept the existence of situated possibilities (I will give a few reasons to accept their existence). I also show a stronger result: for all practical purposes, such relations of necessity are no more and no less justified than a corresponding universal generalisation, so that it would be irrational to accept the latter while rejecting the former. The consequence of this is that empiricist-minded philosophers have no reason not to endorse modal empiricism.

Note that there has been a resurgence in empiricist epistemology for modalities in recent years. Some authors have argued in favour of the idea that knowledge of metaphysical possibilities is attainable by induction on experience (Roca-Royes 2007; Hawke 2011; Strohming 2015). I am sympathetic to these accounts, and mine shares similarities. However, it differs in its focus on natural modalities rather than metaphysical ones, which makes it more modest in this respect. It is also more ambitious in other respects, because I aim to show that one can have knowledge not only of possibilities, but of relations of necessity from experience, and that they are as justified as universal regularities.

Let us start by presenting the kind of situated modalities involved in modal empiricism's notion of empirical adequacy. I will then explain how they can be known by induction, and finish with some remarks on the epistemology and metaphysics of the laws of nature.

5.2 Situated Possibilities

I have introduced the notion of situation in the previous chapter (Sect. 4.2.2). A situation is a local, bounded, occurrent state of affairs that one could represent. We describe a situation using our language, possibly using theoretical language, for example values for measured properties in a scientific context. This allows us to classify situations into *types*, corresponding, for example, to the kinds of objects they contain and to the way they are related to their environment, and to assign a *state* to them, corresponding to definite properties and relations for the objects they contain, including specific values for particular degrees of freedom.

What I call a possible situation is an alternative way a given situation could be. This is the same situation, of the same type, but in a different state. When I say “possible”, I am not talking about epistemic, conceptual nor any kind of mind-dependent possibilities, but about alethic possibilities, that is, mind-independent possibilities “out there, in the world”. To say it differently, whether or not a situation is possible is not known a priori.

Let us first see how this notion differs from other kinds of possibilities, such as metaphysical and nomological possibilities. I will then give a few reasons to accept them in our ontology.

5.2.1 *Conceivable and Possible Situations*

It is common to think of nomological possibilities as a sub-set of conceivable possibilities: only part of what can be conceived is allowed by the laws of nature. Situated possibilities are similar to nomological possibilities in this respect, except for their situatedness (I consider that both nomological and situated possibilities are kinds of *natural* possibilities), and an intuitive way of presenting them is first to consider the set of *conceivable* situations.

Imagine that a scientist wants to test the laws of classical optics. She aims a laser towards a reflecting surface from a particular angle and observes that the light ray is reflected by the surface at the same angle.¹ This is a situation described from a certain perspective: a focus on incident and reflection angles with a certain degree of precision. There is a set of conceivable situations that correspond to it: the ray could have been aimed and reflected at any angle within the range of angles that we could discriminate, all else being equal. This set, which corresponds to various conceivable manipulations and observations of the situation, can be enumerated a priori (in physics, they would correspond to coarse-grained regions of the state-space of the represented system). But because of natural constraints, only some of these situations are naturally possible: as we know from the laws of optics (or as we shall assume for the sake of this illustration), those possible situations are the ones where the light ray is roughly reflected at the same angle as the incident angle. All other conceivable results are naturally impossible.

Conceivable states for targets of representation were already mentioned when introducing the notion of context in Sect. 3.2.1. I also mentioned conceivable contexts, including conceivable manipulations on situations, in Sect. 4.2.2. Combining these two processes gives us a set of conceivable situations. The idea is to imagine the range of states in which an actual situation, to which we refer rigidly, could be while remaining the same situation of the same type with the same objects if it were measured or manipulated differently.² Possible situations can be thought of as a sub-set of this range: the ones that are naturally possible. Since it results from natural constraints, this sub-set is not known a priori: it is not a matter of meaning or conventions, but of real possibilities in the world.

I assume that a situation is identified at least by its type. This type must be preserved when considering alternative states the situation could be in. We must still be talking about a laser pointing at a reflective surface when considering alternative possibilities. This is important, because the range of conceivable situations, hence the range of possible situations, depends on the way the situation is identified. Several situations with different modal characteristics can share a material constitution. To take Aristotle's famous example, there are alternative states a lump of clay could

¹This is simplified for the sake of the illustration. Optical phenomena are more complicated than this.

²As noted in Sect. 4.2.2, norms of experimentation and theoretical frameworks could be involved in this imagination process, but this is not necessarily problematic.

be in which do *not* correspond to alternative states a statue constituted of this lump of clay could be in, because this alternative state for the lump of clay would no more constitute a statue. So, a situation identified as containing the statue is not a situation identified as containing the lump of clay, because the ranges of conceivable states they could be in are different. The same kind of remark can be made assuming that a given coarse-graining is part of the characteristics of a type of situation: a coarse-grained description will not discriminate between some possibilities that a more fine-grained description would, so the range of conceivable situations associated with these types is different.

One could wonder what these types of situations are: do they correspond to natural kinds, or to conventions? In the former case, are we not conceding too much to the realist, and in the latter case, can we maintain that possible situations are objective aspects of the world? My answer would be this: it does not really matter how situations are identified. Their bounds in space and time, for instance, can be delimited arbitrarily: as noted in the previous chapter (Sect. 4.2.2), we could consider the trajectory of a single planet, or of all planets in the solar system, during a few days or a few years, and all of these choices succeed in delimiting situations. However, once their boundaries are delimited, situations have objective, empirically accessible characteristics. The same goes for their type: we could consider any identity condition, for example, functional identifications for organs in biology, or an identification in terms of material constitution in other cases. All of these choices will be associated with broader or more limited ranges of conceivable states, as the statue and clay example shows. But once the choice is made, once the range of conceivable states is delimited, there is an objective fact of the matter about which of these conceivable states are possible and which are not. So, type identification is indeed conventional, and actually, the way we classify situations into types is generally theory-laden (it depends on experimental norms: see Sect. 4.2.2), but this does not mean that possible situations are mind-dependent entities.

Having said that, I presume that some conventions for identifying situations are more fortunate than others, because the associated types are projectible, in the sense of Goodman (1954) (which will translate into their instances having shared modal characteristics). I will say more about this later.

Let me emphasise an important aspect of this picture: possible situations should not be conceived of *in abstracto*, but they are, so to speak, anchored to an actual situation of reference. In the illustration above, when conceiving a different angle at which a light ray could be reflected, we are not talking about any random situation anywhere and at any time, nor about an abstract type of situation, but about *this very situation at this place and time*, albeit in a different state. This is important if we want to say, as suggested in the previous chapter, that our models of proteins are relevant to actual living tissues for instance.

A good way of accounting for this “anchoring” aspect is to borrow Kripke’s principle of necessity of the origin. The idea is that all alternative ways a given situation could be share a common origin, which can be traced back by considering the continuous evolution of the situation from its origin to the present state,

demanding that it remains a situation of the same type. One need not go very far: perhaps tracing back to the moment the situation was first identified by ostentation is enough.³

This rigid reference to an actual situation implies fixed background conditions: these background conditions, and the environment of the situation more generally, are preserved across possibilities. For example, if a situation involving an object in free fall is located on the surface of the Earth, then this is also the case for all alternative ways the situation could be. These background conditions do not characterise types of situations, but only particular instances. They are not in general known exhaustively by someone representing a situation (which will be important later), rather they could be expressed by a *ceteris paribus* clause, pointing at the situation and assuming that some degrees of freedom of the situation could vary, all else remaining equal. So, these background conditions are shared by all alternative ways this particular situation could be. They can affect which among the conceivable states the situation could be in are possible and which are not, because constraints on possibilities can depend on the environment. For example, the possible trajectories of an object in free fall depend on the presence of nearby massive objects.

Note that these background conditions do not entirely determine the set of possible states the situation could be in. By stipulation, the particular state of a situation of a given type is not part of its identity, but a contingency. This allows, for example, to consider various possible manipulations we could make on the target.

To sum up, a situation is characterised by a bounded range of conceivable states, and this range is determined by the way the situation is identified, notably, by its type. This range is somehow conventional, generally theory-laden and known *a priori* (assuming experimental norms). Natural and environmental constraints associated with the type, origin and background conditions of the situation further limit this range to a set of possible situations, which is not known *a priori*.

Because of all the characteristics mentioned above, possible situations are different from possible worlds. More precisely, they differ from possible worlds in two important respects:

- Extensional limitation: Possible situations are not maximal states of affairs. They are situated, bounded in space and time and contain a finite set of objects.
- Intensional limitation: The range of possibilities for a situation is also limited in a way possible worlds are not. The identity of the situation, in particular its type, origin and background conditions, must be preserved when considering alternative states a situation could be in.

³Note that I take this principle to be a useful convention rather than a metaphysical principle, since situations are identified conventionally in any case. Maybe other conventions could be considered. In any case, the necessity of origin will not play an important role in the arguments of this book, but the anchoring aspect will.

Modalities are often represented using Kripke's possible world semantics, by introducing an accessibility relation between possible worlds. In the $S5$ system, this relation is reflexive, symmetric and transitive, which means that all possible worlds are accessible from all others. Similarly, situated possibilities can be represented by considering an accessibility relation between actual situations and their alternatives, using $S5$ modal systems centred on every actual situation, each limited to the relevant objects in this situation. The situations that are accessible from a given actual situation are the possible ways it could be. This semantics provides an interpretation for the modal operator $\Box_S$ that was used in the previous chapter: this operator quantifies over all possible situations that are accessible from a given actual situation S .

In this chapter, I assume that possible situations exist (I give a few reasons below). I do not want this assumption to be too metaphysically loaded: I do not claim that possible situations are concrete entities, or that they represent dispositions or relations between universals. All I assume is that there are alternative ways actual situations could be, and alternative ways they could not be in virtue of natural and environmental constraints. I assume that it is a matter of empirical inquiry to know which, among the conceivable situations, are naturally possible and which are not, and the purpose of this chapter is to show that such knowledge is attainable by induction on realised possibilities. Such knowledge will take the form of necessary statements that characterise all the possible situations of a certain type, and these alone.

Let us say a bit more about these statements, and the kind of necessity involved.

5.2.2 *Statements of Necessity*

In the framework introduced in Chap. 3, sets of conceivable, coarse-grained states for situations were represented by means of contexts. The set of conceivable states specified by a context is limited by a choice of relevant variables for an agent (in particular a distinction between fixed and variable properties: see Sect. 3.2.1). As noted earlier, considering all the contexts afforded by a situation and its *conceivable* alternatives (assuming its type and origin are preserved) give us a set of conceivable states for this situation.

We have seen that interpreted models bring restrictions on the set of states specified by a context: only some of them are “permitted” by the model, and the model is accurate if the actual state of the target is among them. A general model is characterised by a function from contexts to sets of licensed interpreted models, so in effect, a general model also brings restrictions on the set of states specified by the set of conceivable contexts in which it has a licensed interpretation. To take an example, the general mechanical model of the pendulum has a relevance function that gives its domain of application, in this case, pendulums, which corresponds to the set of contexts for which it has licensed interpretations. We can associate this domain with the definition of a *type of situation*, which is the type represented

by the general model. This type is characterised by the features shared by all contexts in which the model has licensed interpretations: the general characteristics of pendulums from a mechanical perspective.⁴ The general model brings restrictions to the possible trajectories of various pendulums, whatever our perspective is on these pendulums (that is, whatever degree of precision, etc.).

It follows that claiming that a general model is modally adequate, taking the definition of adequacy of Sect. 4.2.6, amounts to claiming that all situations of a certain type are necessarily in a certain way. The type of situations involved characterises, as just said, the domain of application of the model. What the general model “says” about these situations corresponds to the conditions of accuracy that its licensed interpretations specify. For the model to be modally adequate, these conditions must be fulfilled in all possible contexts and situations in its domain. So the claim that a general model is modally adequate is a claim of the form: $(\forall s)As \rightarrow \Box\Phi(s)$, where A denotes a type of situation.

The reverse is likely to be true as well: any statement of necessity of the form $(\forall s)As \rightarrow \Box\Phi(s)$ can be turned into a statement to the effect that a general model, whose relevance conditions are limited to objects of type A , is modally adequate.

Take, for example, the statement that all swans are white as a matter of necessity. This can be translated into a statement about a very simple general model. This model has two symbols: one object, s , and one property, w , and s is in the extension of w . The relevance function of the model is the following: (1) the model has licensed interpretations only in contexts where the target of representation is a swan and where the colour of the target is the property of interest, and (2) in these contexts, object s is mapped to the swan and property w to the property of being white. What the model does, in these contexts, is limit the set of permitted states for targets to the ones corresponding to being white. The model is modally adequate if all its licensed interpretations are accurate for all possible situations in the universe, which is to say that all possible swans are white (that is, all alternative ways swans could be are such that they are white).

So, at least simple statements of the form “necessarily, all swans are white” can be translated into the statement that a model is modally adequate. The same rationale could be given for universal generalisations and non-modal versions of adequacy. I presume that scientific models have all the resources necessary to express more complex statements: for example, that all emissions or absorptions by hydrogen atoms should have a frequency corresponding to one of the lines of the spectrum predicted by the model of the hydrogen atom, or that all trajectories of solid bodies around a black hole should respect the constraints of a particular model of relativity theory. In the latter case, the same model can apply to situations of the same type (black holes) even if the mass of the black hole is distinct in each case. The kind of statements that might not be translatable into models are statements that do not concern objects of a certain type, but perhaps the universe in general, or abstract

⁴This type is perspectival, since it is primarily characterised by a set of contexts. It can be conceived of as a type of activity. See Sect. 3.3.3.

entities. Such statements are more metaphysical in nature, and I am not interested in them here.

What I am interested in is the kind of necessity involved if necessity statements are interpreted in this way (which is a kind of *de re* necessity about situations). These statements are not made true by possible worlds, but by alternative ways actual situations could be. As mentioned above, this implies intensional limitations to the range of possibilities considered, when compared with possible world semantics. Let us examine them.

Some of these limitations can be classified as a type of coarse-graining. The general idea is that we are blind to some counterfactual variations, either because they are not the focus of the statement considered or the associated model (a mechanical model of pendulum describes properties such as position and velocity, but remains silent about the colour of the pendulum, for instance) or because they are beyond our discrimination abilities (remember that an interpretation involves accessible properties only). For example, variations in the size of the swan do not matter; only the colour does. The term “white” does not denote a very precise property, but variations in whiteness do not matter, or perhaps cannot be discriminated. Counterfactual situations where the swan would not be a swan anymore are *de facto* excluded. All these limitations will translate into the fact that our general model has no licensed interpretation in some contexts. This kind of limitation is typically associated with the vocabulary used by the model (here, “swan” and “white”), in the same way the information brought by a city map is limited by the symbols it uses. A certain notion of relativity is implied, which is a relativity to a class of purposes associated with particular domains of experience.

These limitations entail that the statements of necessity we are concerned with are closer to phenomenological laws, or what Duhem (1906) called “experimental laws”, than they are to fundamental laws of nature. Assuming that there are such things as laws of nature, such statements of necessity would at best correspond to approximations of these laws restricted to particular types of objects, activities and accessible properties. Take, for example, Bohr’s model of the hydrogen atom. This model predicts the frequencies of emissions and absorptions of light by hydrogen atoms. Limiting ourselves to this particular prediction for the sake of the illustration, claiming that this model and this model alone is empirically adequate is tantamount to claiming that Balmer’s formula (or one of its generalisations), an experimental law predicting spectral lines, is correct. Perhaps the model can predict more than this, but by no means does the empirical adequacy of Bohr’s model entail that Schrödinger’s equation is correct, for instance, and the same remarks could be made with respect to a relativistic model of black holes and relativity theory (this is because any model incorporates domain-specific posits: see Cartwright (1983, ch. 6)).

There is another kind of limitation that is more specific to the situatedness of possibilities of the present framework. It has to do with the fact that the possibilities considered are anchored to actual situations. When claiming that all swans are necessarily white, we are considering alternative ways that actual, situated swans could be, and not merely possible swans. The identity of these swans, in particular

their origin, is considered fixed. Even considering a situation that comprises the entire life of a swan, from its birth to its death, we are still rigidly referring to this particular swan with particular genes, assuming that these genes are determined by the origin of the swan.

The range of actual situations to which we have access could be limited by our position in the universe. Some situations could be out of reach in principle (for example, situations inside black holes), and if the possibilities that matter for the justification of a statement of necessity must be anchored to actual situations that are *in principle accessible*, then the associated contexts are not among the ones we can consider. These limitations entail that the necessity involved here is markedly weaker than the necessity usually associated with the laws of nature (hence weaker than metaphysical necessity).

To show this, assume that there are such things as laws of nature, and imagine a possible universe with the same laws of nature as our universe, wherein a civilisation is interested in the acceleration rate of free-falling objects. However, in their universe, all objects are situated near the surface of the Earth. The laws of nature of this world, which are the same as ours, do not preclude an object free-falling at an acceleration rate higher or lower than 9.8 m/s^2 , but in their universe, it is still true as a matter of necessity that all objects fall at this rate, because this is true of all alternative falls of actual objects, given that these objects are all situated near the surface of the Earth, and that background conditions are considered fixed for all alternative ways these objects could fall. Note that the same would go if instead the limitation was epistemic: if, for any reason of principle, the civilisation had no access to situations far from the surface of the Earth. But the statement “necessarily, all objects fall with an acceleration rate of 9.8 m/s^2 ”, although true of all situations to which this civilisation has access in principle, is hardly a law of nature, let alone an approximation of a law, because it depends on contingencies associated with the situation of this civilisation.

This example shows that the kind of necessity associated with situated possibilities is relative to our epistemic situation in the same way that technical modalities (as in “it is not possible to travel from Brussels to Hong Kong in one hour”) are relative to our technical abilities. This means that this notion of necessity is more permissive than nomological necessity: more relations count as necessary in this sense. On the other hand, this means fewer possibilities. One could construe this kind of necessity as expressing not natural constraints directly, but the way these natural constraints express themselves in the contexts to which we have access, given background conditions shared by these contexts. Furthermore, as observed earlier, only a certain type of context is concerned by a statement. Statements of necessity are thus relative to a perspective associated with our epistemic position in the universe, and with a focus on a certain type of situation.

This is the kind of necessity involved when a modal empiricist asserts that a theoretical model is empirically adequate, and I will argue that this kind of necessity is knowable. My aim is not to show that we could have access to metaphysical essences, let alone to the fundamental “modal structure of the world”, as structural realists would have it (French 2014; Ladyman and Ross 2007). The

modal structure of our possible interactions with reality would be more appropriate, since a perspectival component is implicit. However, in contrast with the way modalities are generally considered by empiricists, I am talking about objective, alethic possibilities: not mere heuristic or fictional devices, not known *a priori*, not degrees of credence or other mind-dependent entities, but possibilities *in the world*.

5.2.3 *Why Should We Accept Situated Possibilities?*

The purpose of the following sections of this chapter is to show that if situated possibilities of the kind presented above exist in nature, then we can have knowledge of them, knowledge which takes the form of statements of necessity. This justifies adopting a modal version of empirical adequacy as the aim of science, because this aim is achievable in principle. However, the arguments that I will propose are conditioned on the existence of situated possibilities. Arguably, the assumption that these possibilities exist at all, an assumption which has metaphysical overtones, should be justified independently. The main reasons to accept it are semantic, pragmatic and empirical, and these reasons are empiricist-friendly. Let us briefly review them.

On the semantic level, modal discourse is ubiquitous in natural languages, and in science as well, and denying meaningfulness to whole parts of our discourse looks like a dogmatic position. This is particularly true for empiricists, like van Fraassen (1980, ch. 1.2), who advocate a “literal interpretation” of scientific theories, since modalities are involved in the motivations for a “literal interpretation”

As explained in Sect. 1.2, the idea that theories should be interpreted literally opposes what Psillos (1999) calls reductive empiricism, which attempts to reinterpret the content of scientific theories in terms of mere observables by means of correspondence rules. This is a version of semantic reductionism. There are many arguments against reductive empiricism. Some have to do with the problematic status of correspondence rules, and the distinction between theoretical and observational vocabulary. They are addressed by a focus on models instead of statements (see Sect. 2.2). Others have to do with the fact that most theoretical terms are dispositional, and that the meaning of dispositional terms cannot be reduced to logical relations between observations because of their modal aspect: a disposition, such as being soluble, can be assigned to an object, such as a sugar cube, without ever being manifested, for example if the sugar cube is never put into water. This problem seems to call for a literal interpretation of *modalities*.

Another type of argument against semantic reductionism is directed against descriptive theories of meaning more generally. According to Kripke (1980) and Putnam (1975), the meaning of a term like “gold” or “water” cannot be elucidated in terms of descriptions, such as “yellow metal” or “transparent liquid” that would hold as a matter of *a priori* necessity. Kripke’s arguments, in particular, rest on an analysis of modal discourse: he argues convincingly that necessity and apriority should be kept distinct. It follows that what holds necessarily of a substance

such as gold can be discovered by experience, and we should not think that the descriptions that we associate *a priori* with such a term, such as being a yellow metal, are a matter of necessity. Putnam notes, in a similar vein, that any description associated with a theoretical term is revisable in principle (such as “being a fish” for whales). The upshot of these analyses is that the meaning of theoretical terms should be understood in terms of direct reference to a kind which typically *causes* the manifestations that we associate with this term, rather than being analysed in terms of these manifestations. Being transparent is not part of the essence of water, but rather the essence of water causes transparency and other superficial manifestations in normal contexts. Note that the concept of causation is modally loaded, and that these problems faced by descriptivism are not very different from the problems related to the reduction of dispositional terms.

My aim here is not to defend a particular kind of semantics, for example, direct reference to natural kinds, nor any version of essentialism (I will suggest adopting a pragmatist semantics in Chap. 8). However, these arguments show, conclusively in my opinion, that modal discourse is distinct from discourse about meaning and analyticity. I presume that advocates of a literal interpretation of theories accept as much, since this is part of the motivations for their view. This conclusion holds whether or not there are natural possibilities in the world, and one should not feel obliged to draw ontological conclusions from semantic arguments, of course. Perhaps modal statements have no mind-independent truth conditions at all. Perhaps they are really meaningless. Or perhaps they are all false and we are all confused when using counterfactual talk. But I would say that since Kripke and Putnam’s arguments, it has become the modal sceptic’s burden to convince us that this is the case. And given the centrality of modally loaded concepts in science, such as explanation and causation, this burden is quite heavy.

Situated modalities are well-suited to account for modal discourse, since in general, counterfactual statements are not maximal and are, so to speak, “anchored” to particular objects to which we refer directly (this can be understood in terms of the notion of rigid reference employed by Kripke, in which a term refers to the same object in all possible worlds). In this respect, my focus on possible situations rather than possible worlds shares some motivations with situation semantics in philosophy of language (Barwise and Perry 1983; Kratzer 2019). Their approach notably emphasises the fact that universal statements whose truth conditions do not depend on the context of utterance are the exception rather than the rule in natural languages. Counterfactual statements generally assume a fixed background, which is contextual. This is the case, for example, with explanations (see Sect. 3.3.6). The explanation of a fire can take the form of a counterfactual: the fire would not have occurred if not for this spark (so the spark explains the fire). In this explanation, the fact that there is oxygen in the air is not considered a relevant explanatory factor, but a background condition. Explanations depend on variables of interest. All these aspects of counterfactual talk are easily captured by our understanding of situations.

Let us now turn to pragmatic reasons to endorse modalities. Many of these reasons have already been mentioned in Chap. 4 (Sect. 4.4.3). Accepting modalities can help make sense of scientific practice as a rational activity. Scientists often

implement experimental situations that would not occur if it were not for their interventions. They control target systems or they vary parameters as if they were exploring various possibilities. From the beginning of modern science, this is an important principle of scientific methodology, which is particularly well expressed by Galileo, when he claims to be “trying to investigate what would happen to moveables very diverse in weight, in a medium quite devoid of resistance”. He then proposes to “observe what happens in the thinnest and least resistant media, comparing this with what happens in others less thin and more resistant” (cited in McMullin 1985, p. 267).

As argued in Chap. 4, these experimental activities make sense if the scientists’ motivation is to know *what would happen* in such situations, as acknowledged explicitly by Galileo, implying that gaining modal knowledge must be part of the aim of science.

Situated possibilities are also particularly well-suited to account for this pragmatic aspect in a minimal way, because the situations scientists implement are not maximal states of affairs, and because idealisations, as employed by Galileo, involve focusing on properties of interest and neglecting others, which can be addressed by the perspectival aspect of situated possibilities that is involved in the identification of situations.

A final class of reasons for accepting modalities are empirical. These reasons come from empirical research on action-based perception.

On action-based views of perception, what agents perceive at particular moments is partly constituted by what they are doing, or what they are able to do (see Briscoe and Grush (2015) for a review).⁵ This can be translated into the idea that we perceive (veridically or not) possibilities for action, or what Gibson (1986) calls affordances: for example, that a door knob is graspable, or that a surface is walkable. Such views are supported by philosophical arguments. For example, the fact that we instinctively attempt to catch a ball thrown at us from behind a plexiglass, even though we know that it is not actually catchable, seems to show that we perceive it as catchable (Nanay 2012, pp. 438–439). It is also supported by research in cognitive science (for example, Prinz 1997; Hommel et al. 2001; Hommel 2004; Schütz-Bosbach and Prinz 2007), which shows that sometimes perceptual representations include a motor component. This research suggests that action plans and perceptual items share a common representational format.

In the context of a naturalised epistemology, in particular its methodological version (Goldman 1994), one could expect empirical science to inform epistemology. More specifically, an empiricist assuming that the notion of observation plays an important role in the justification of scientific theories could turn to cognitive science in order to inquire about this notion. According to action-based views of perception, the conclusion would be that our observations are strongly associated with modal characteristics. In particular, possibilities for action are sometimes part

⁵I am indebted to Andrew Sims for my acquaintance with these accounts, as well as for the technical details that follow.

of the content of perception. Note that the ability to correctly perceive possibilities for action can be learnt by experience, but this is true of perception in general: one can learn how to recognise birds or letters, for instance. The way perceptual expertise about possibilities could be achieved is actually compatible with the inductive epistemology that will be presented in this chapter.

Just as in the case of semantic arguments, one should not feel obliged to draw metaphysical conclusions from these empirical observations. Perhaps the modal component of perception when one perceives a ball as catchable is illusory, for instance, and not only in particular cases (no one denies that perception is fallible): perhaps it is a massive illusion, because there are no possibilities in the world. However, the most natural move consists in taking this perceptual content at face value, that is, accept that there are possibilities in the world which we can perceive, and (assuming action-based views) the burden is on the sceptic to show why we should think that this is a massive illusion.

Again, the situatedness of the kind of modalities considered here has good compatibility with action-based accounts of perception, since possibilities for action are anchored to actual objects and relative to the perspective of an agent, in particular its abilities. As we have seen previously, part of the motivation to incorporate modalities in our notion of empirical adequacy stems from the interventionist aspects of experimentation, which is also in line with these accounts of perception. The idea could be that scientific knowledge is nothing but an extension or a sophistication of mundane modal knowledge, assuming a wider range of possibilities for action.

So, we have *prima facie* good reasons to accept possibilities in our ontology, that is, that there is a fact of the matter about whether the situations we encounter *could have been different* in one way or another, all else being equal, in virtue of natural constraints. As far as I can see, a principled impossibility for modal knowledge is the main reason to assume that modal statements have no truth-value, but the aim of this chapter is to show that this reason is not tenable in the case of situated modalities. So, let us now explain how modal knowledge is attainable.

5.3 The Inductive Route Towards Necessity

In the previous section, we have examined the notion of situated possibility that characterises modal empiricism. We have seen that statements of necessity, if interpreted in terms of possible situations, are weaker than statements interpreted in terms of nomological or metaphysical necessity, in particular because they are relative to our epistemic position in the universe. The claim that a model is empirically adequate is, per the definition of Sect. 4.2.6, a statement of necessity of this kind.

My aim in this section is to show how such statements of necessity (and so, empirical adequacy) can be justified by means of an “empiricist-friendly” type of inference: induction. This is a mode of inference by which one starts from the observation of a regularity in a sample of exemplars of a given kind, and infers

that this regularity holds for all entities of the same kind. In our case, the exemplars will be realised possibilities, and the larger set the set of all possibilities.

Let me start with some generalities about the justification of induction.

5.3.1 *Induction and Underdetermination*

In this work, I assume that induction is a valid mode of inference, in the following weak sense: given that a sample of experienced situations of a given type displays some regularity, we have *prima facie* good reasons to believe that this regularity can be extended (or “projected”) to the set of all situations of the same type, so long as no relevant conflicting statement that we have knowledge of is equally justified. Hence it is often reasonable to believe statements of universal regularity. I do not claim that such beliefs cannot be defeated by more evidence, only that it is reasonable to believe them as long as no evidence to the contrary is at our disposal, and as long as they are not in competition with other equally well-justified statements.

I cast general statements in terms of situations. As explained in the previous section, it is likely that this can be done without loss of generality, assuming that a statement like “all swans are white” can be translated into “all situations where there is one swan are situations where there is one white swan” (or the claim that a corresponding model is adequate).

The clause about competing statements is important because in principle, we could have conflicting statements that are both justified by induction. For example, an induction on gold spheres could tell us that they are stable whatever their size, and an induction on metal spheres in general could tell us that they become unstable past a given size. Since gold is a metal, these two results are in conflict. They are underdetermined by evidence. I would say that we should suspend our judgment in such a case and wait for more evidence. But I would also say that there is no such problem in the case of statements that have no known serious competitors.

It might be hard to tell what counts as a serious competitor. One might worry that there is always a possibility of coming up with an inductively justified statement in conflict with any statement, even if it is far-fetched. This problem is somehow related to Goodman (1954)’s new riddle of induction: why would the predicate *blite*, meaning “white before *t* and black after *t*” not be valid for inductive reasoning? In Goodman’s terminology, why would it not be “projectible”? If it was, then “all swans are white” would be as justified as the incompatible “all swans are *blite*”.

There are other sceptical arguments against induction, tracing back to Hume at least, according to which induction rests on the postulate that nature is uniform, which cannot be justified by deduction (because we need an ampliative inference) nor by induction (because it would be circular). Concerns about induction can in general be reframed as problems of underdetermination by available evidence. Hume’s scepticism, for example, could be framed, taking the case of swans, as resulting from an underdetermination between the statements “all swans are white” and “all swans are white until now and non-white from now”: before the observation

of a black swan, both are compatible with the same available evidence, so we cannot have any justification for one or the other without adding presuppositions.

This issue has been turned by some authors into a reduction against the kind of sceptic epistemology that would warrant induction but refrain from accepting more ambitious modes of inference, such as inference to the best explanation: if all ampliative modes of inference suffer from a problem of underdetermination, why set the bar at one height rather than another? Why not be more ambitious? This has been mounted as a defence of modal knowledge, assuming that such knowledge is arrived at by inference to the best explanation (Ladyman and Ross 2007, ch. 2.3). It has also been argued that induction ultimately rests on an inference to the best explanation (I return to these arguments in Sect. 5.4).

I will not address these problems in depth in this chapter (see Williams 1963; Stove 1986; Campbell 2001; Wright 2018, for defences of induction). I will simply assume that some predicates are projectible while others are not, and that therefore, some statements can be justified by induction, while others, such as “all swans are blite”, cannot. I presume that most predicates of ordinary language are projectible, and that the projectibility of theoretical predicates is ensured by the norms of experimentation which connect theoretical and ordinary predicates. It is ensured, in particular, by the fact that these norms are directed towards the stability and reproducibility of experimental results (see Sect. 3.4.2).⁶ I will talk about underdetermination only when two justified statements are in conflict.

Like Ladyman and Ross, my focus is on modalities, but my defence of modal knowledge will take a different route. I do not wish to argue that modal sceptics place the bar at a wrong position. I wish to argue that at least for situated possibilities, wherever one places the bar in these matters, universal generalisations and relations of necessity will fall on the same side of it: given a universal generalisation of the type “all A are B” and its modal counterpart “all A are necessarily B”, either both are underdetermined by evidence, or neither are. So, there is no reason for someone accepting induction towards universal generalisations not to accept the possibility of modal knowledge.

This seems counterintuitive, because modal statements are strictly stronger than their non-modal counterparts. My main argument is that although one statement is indeed stronger, this cannot make any pragmatic difference concerning what it is rational to believe or not, so that for all practical purposes, they will always fall on the same side of underdetermination (or to be precise, one cannot be in a position to *know* that one of the statements is underdetermined but not the other). This follows at least as far as situated possibilities are concerned.

⁶This idea is not very different from Goodman’s proposed solution, that projectible predicates are “entrenched”.

5.3.2 *The Inductive Route*

So much for the preliminaries. Let us now turn to the main question of this section: can at least some relations of necessity be justified by induction?

The received view, as far as I know, is that they cannot, whatever the kind of alethic necessity considered. Here is a short rationale that bolsters this view. Induction is an ampliative mode of inference by which one infers, from a sample of exemplars of a given type displaying a regularity, that this regularity holds for a larger set: the set of all objects of this type. Hence, what can be known by induction is a universal regularity. But a relation of necessity goes beyond that. It does not only concern the actual world, but all possible worlds. However, we have no epistemic access to other possible worlds (no “modal telescope”); we do not experience relations of necessity or mere possibilities. Possible worlds, if they exist as such, are causally disconnected from the actual world, so we cannot observe their effects either. An induction on possible worlds will not do: with only one exemplar at our disposal, the actual world, the induction cannot get off the ground.

If one accepts this rationale, there are two options.

The first option is to be a modal sceptic and assume, for example, that modal talk is a mere pragmatic way of speaking that does not purport to be objectively true or false (van Fraassen 1989), or perhaps that modal talk is meaningful, but generally false. Taking the case of nomological necessity, one can deny that there are such things as laws of nature (which is van Fraassen’s view), or one can assume that the laws of nature supervene on non-modal facts, as in the best system approach (Lewis 1973, pp. 72–77).⁷ Then the laws of nature might be knowable by induction, but they have no “modal force”: they merely summarise the regularities of the actual world. I have given in the previous section some reasons to reject this modal-sceptic option.

The second option is to accept that there are other modes of inference than induction, typically, inference to the best explanation. As far as I know, this is the option followed by all philosophers who assume that the laws of nature have a “modal force”, and do not supervene on non-modal facts: they assume that these laws play an explanatory role, and that they are justified as such. The rationale can be the following. One sees some regularity in a sample of phenomena: all *F*s in the sample are *G*s. The best explanation for this observed regularity is that as a matter of nomological necessity (that is, as a consequence of the laws of nature), all possible *F*s are *G*s. From this one can infer that all actual, past, present and future *F*s are *G*s, but one did not arrive at this result by induction: it was arrived at by making a detour through inference to the best explanation (Dretske 1977; Armstrong 1983; Foster 1982). I will present and reject some motivations for this approach in Sect. 5.4.

⁷Although, as is well known, Lewis was not a modal sceptic.

The problem with this second option is its use of inference to the best explanation as a principle of *justification*. Empiricists typically deny that the non-empirical criteria that are constitutive of a good explanation, such as simplicity, are truth-conducive. They could play a pragmatic or heuristic role when it comes to proposing new hypotheses, but from an empiricist perspective, this tells us nothing about their validity: these hypotheses should then be put to the test so as to ensure that they are empirically adequate, and the justification of their empirical adequacy can, at best, be inductive (I will return to this theme in the next chapter).

Our two options are therefore problematic. However, I would say that we are facing a false dilemma: we do not have to choose between the truth-conduciveness of inference to the best explanation and modal scepticism. The received view is wrong. One can justify necessity by induction alone, at least when it comes to the situated necessity considered here.

This is because the possibilities we are considering, contrarily to nomological or metaphysical possibilities, are situated: one should think of them not in terms of possible worlds, but in terms of possible situations. Then all that is required to justify a claim of necessity by induction is to assume that the situations we experience are a representative sample of the larger set of possible situations of the same type. As an example, one can take the observation that all objects dropped at the surface of the Earth accelerate towards the ground at a certain rate and infer that this is true of all possible situations of the same type, which is to say that it is true as a matter of necessity, at least relative to some background conditions that happen to be present in our surrounding environment. This means that one is in a position to assign truth-values to counterfactual statements such as “if I dropped this object, it would fall towards the Earth”.

An induction on possible worlds would of course be problematic, since our sample would consist in only one exemplar. If possible situations are construed as alternatives to actual ones, an induction on the alternatives to a single situation would be problematic as well, for the very same reason. But this is not so with an induction on all possible alternatives to all situations, at all places and times in the universe, assuming that observed situations are representative not only of all other unobserved actual situations of the same type, but also of all unobserved alternatives to all other situations of the same type in the universe. In this sense, the actual fall of a dropped object can be considered representative of the counterfactual fall of an object that was not actually dropped (assuming this object *could have been dropped*, which can also be known by induction on objects of this type). Similarly, the various manipulations performed by scientists in experimental contexts can be considered representative. For example, the manipulations involved in observing the characteristics of proteins mentioned in Sect. 4.2.4 are representative of what *could have been observed* on other samples of living tissue. This kind of induction is defeasible. New evidence might prove a relation of necessity wrong. But the same goes for any kind of induction.

Note that this account is not very different in spirit from Roca-Royes (2007)’s similarity-based epistemology for *de re* modalities. She argues that one can have knowledge of unrealised possibilities on the basis of realised possibilities: for

example, the fact that a table breaks indicates that it is possible that it breaks (because actuality implies possibility), and this modal knowledge can then be extended to other similar tables, to the conclusion that they are breakable.

This approach can be qualified, in Hale (2003)'s terminology, as a possibility-first approach: the idea is that knowledge of a given class of possibilities, and in particular, the absence of some relevant possibilities in this class, informs us about relations of necessity. Here, the relevant class of possibilities is the class of realised possibilities.⁸

The assumption that observed situations are representative of possible situations cannot itself be justified by induction without circularity. But this is no more a problem than it is for other versions of induction, where one assumes that observed situations are representative of all unobserved ones of the same type in the universe. The notion of representativeness involved here is, of course, sensitive to the way situations are classified into types. One could wonder at which point some variation in an experimental setting makes the type of situation different, and not representative anymore. I have claimed earlier that types of situations can be defined conventionally (Sect. 5.2.1), but that some choices are more fortunate than others. This is also related to the notion of projectibility mentioned in Sect. 5.3.2: a property attributed to situations of a type is projectible in so far as these situations are representative of other situations of their type with regard to this property. As explained earlier, I presume that the assumption that a theoretical predicate is projectible rests on experimental stability and reproducibility, so this is itself known by induction (assuming that more mundane predicates associated with our observations are projectible), and the same goes for representativeness. Any assumption of representativeness is, of course, defeasible. In any case, the main point that I wish to make here is that this kind of problem is not specific to modal knowledge: small variations in experimental designs can also be observed among the set of actual situations. Yet, scientists generally consider that the phenomena observed in such designs are representative of their type.

As we can see, assuming situated possibilities, there is a homogeneity between an induction towards universal regularities and an induction towards necessity: both involve generalising from a sample of situations to a larger set of the same type. One could reject the latter while accepting the former assuming some disparities between the two cases: either merely possible situations do not exist while unobserved actual situations do, or observed situations are representative of unobserved actual ones, but not of merely possible ones. I assume that possible situations do exist (qua possible states of affairs), and in this context, claiming that observed situations are representative only of actual ones seems to be question-begging. Actual situations are merely a sub-set of the possible situations: the ones that are realised. There seems to be no more reason to assume a bias in the way these are selected among

⁸Note that Hale explicitly rejects this option, but (1) he does not discuss inductive reasoning in his paper, and (2) he is mainly concerned with knowledge of metaphysical modalities, and assumes, wrongly in my opinion, that knowledge of weaker modalities must derive from it.

the larger set of possible situations than there is to assume a bias in the way actually observed situations are selected among the larger set of observable ones.

The fact that, in this view, knowledge of necessity is arrived at by induction allows one to defuse a lot of arguments against the possibility of modal knowledge. Take, for example, Hume's contention that no necessary connections are found in experience. No universal regularities are found in experience either; yet universal regularities can be justified by induction, so in our case, Hume's argument against modal knowledge boils down to his argument against induction in general.

Another contention cited above is that possible worlds are causally disconnected from the actual world: we have no "modal telescope" to observe mere possibilities. However, situated possibilities are not disconnected from actual situations in the same way possible worlds are. Arguably, possible situations are among the possible effects of actual ones. For this reason, we possess a sure way to explore possible *types* of situations: all we have to do is implement them through controlled interventions. If we want to know what *would* happen if a laser were aimed at a reflecting surface at an angle of precisely 32°, all we have to do is to implement the situation. It will be representative of all exemplars of its type.

One could argue that once the situation is implemented, it becomes actual rather than merely possible, and more generally, that all situations that we will ever experience are actual, so that we never have knowledge of mere possibilities. Note that we have seen reasons to doubt this: mere possibilities could be part of the content of perception (Sect. 5.2.3). But more importantly, the same problem affects an extension to all actual situations: all situations we will ever experience are observed, so we have no knowledge of unobserved ones. Yet, we are generally willing to extend our knowledge of observed phenomena to unobserved, but observable phenomena. This is nothing but ampliative reasoning. It seems legitimate because unobserved or merely possible situations are of the same type (arguably, an inference to the unobservable is more problematic). So again, this type of argument affects any version of induction.

Finally, relations of necessity have been considered illegitimate by some empiricists because they are posited as explanations of empirical regularities, and inference to the best explanation, as we have seen, is generally rejected by empiricists (van Fraassen 1989, ch. 4.7). This is not the case here: relations of necessity are not explanations, but mere extensions of regularities to the realm of possibilities. They are not arrived at by inference to the best explanation, but by induction. Relations of necessity supervene on the mosaic of possible facts in the same way universal generalisations supervene on the mosaic of actual facts.

In sum, assuming that there are situated possibilities in the world, statements of necessity are in the same boat as universal generalisations with regard to epistemic justification. There is no principled obstacle to modal knowledge from the perspective of someone who accepts the validity of induction. This means that modal empiricism is an acceptable position, even for an empiricist who is sceptical about the "inflationary metaphysics" that can only be justified by inference to the best explanation.

5.3.3 *Modal Underdetermination*

So far, I have argued that knowledge of situated necessity can be arrived at by induction. This means that the modal empirical adequacy of a model, as defined in Sect. 4.2.6, can be justified by an induction on particular uses of this model. However, one might remain a modal sceptic on the grounds that these relations of necessity are strictly stronger than universal generalisation, by adopting a principle of metaphysical parsimony. Perhaps one could adopt van Fraassen's epistemological voluntarism, according to which it is not irrational to have strong realist commitments, but neither is it to have more modest ones, and this is the stance that the empiricist adopts. Accordingly, one could say that it is not irrational to be a modal empiricist and to accept an induction on possibilities, but that it is not irrational not to do so. Claims of necessity are *prima facie* stronger than universal generalisations, and metaphysical parsimony could enjoin us to remain uncommitted to the former while endorsing the latter. This could take the form of an argument from underdetermination: statements of necessity would be underdetermined by actual phenomena, while universal generalisations are not, so it is more reasonable to only accept the latter, and to adopt a version of empirical adequacy that only quantifies over actual situations (see Sect. 4.2.3 for the different versions of empirical adequacy).

I wish to argue, against this idea, that it is actually irrational to refrain from accepting a modal version of empirical adequacy once a non-modal version is accepted, because no practical underdetermination could affect the former without thereby affecting the latter. This means that both are as warranted, from the same evidence and with the same inferential standards (at least once one accepts that there are situated possibilities). The argument generalises to any kind of statement: once a universal generalisation is accepted, there is no reason not to endorse the corresponding statement of necessity.

As explained before, my focus here is on the underdetermination of competing statements. Let us say that two statements are underdetermined if they are *incompatible* (they cannot be both true), but no observation can help us decide which one is true. Translated into talk of models, two models are incompatible if there are possible situations and contexts where the states "permitted" by the two models are mutually exclusive, and they are underdetermined if we are in no position to know which one is accurate in these situations and contexts.

Note that different ranges of relevant observations from which statements or associated models are evaluated, for example, only past observations, or all past, present and future observations, or all possible observations of actual phenomena, will determine different kinds of underdetermination: one can set the bar at various positions. A version of underdetermination that is often considered in the literature is an underdetermination between theories that have exactly the same empirical consequences (it will be examined in the next chapter). This is an underdetermination by all *possible* evidence, and obviously, the modal version of empirical adequacy is no more underdetermined than a non-modal version in this sense, because possible

evidence could help us tell which of two theories are empirically adequate for both versions of adequacy. Other versions of underdetermination could give different results. However, the argument of this section does not depend on a fixed choice in these matters, the point being that any choice will result in necessary statements and universal generalisations falling on the same side of it, for all practical purposes.

I consider that the statements that can be justified by induction are perfectly general: they do not refer to specific objects, and they are not restricted to particular places or times (in logical parlance, they contain no proper names). They concern all situations of a certain type in the universe. This is the case of statements claiming that a general model is empirically adequate in so far as the domain of application of this general model is not restricted extensionally. As an example, a model of optics where a ray of light aimed at reflecting surfaces from an angle is reflected at the same angle is general in this sense, because it can be interpreted in terms of any reflecting surface and any ray of light at any place and time in the world. Presumably, most scientific models respect this condition.⁹

Consider a general statement r and its modal counterpart m . As an example, r could be the statement that all actual gold spheres of more than 1 km in diameter in the universe are stable, and m the statement that all *possible* gold spheres of more than 1 km in diameter are stable. Both statements can be expressed as the claim that a particular model of gold spheres establishing a relation between their size and stability, call it M , is empirically adequate, using a non-modal version of empirical adequacy for r and a modal version for m .

Arguably, m is strictly stronger than r because it quantifies over all possible situations. It would seem, then, that m could be underdetermined by evidence when r is not. Let us examine what it would look like.

Concretely, this means that m could have a competitor that is as well-justified inductively, call it m' , so that it is impossible to decide which of m and m' is true on empirical grounds. m' could be, for example, the statement that all possible metal spheres of more than 1 km in diameter are unstable. This statement could be associated with a model of metal spheres, call it M' , which would imply “permitted states” incompatible with the ones permitted by M in at least one possible situation (one with big gold spheres). m' would be justified by observations of metal spheres (excluding gold spheres), while the incompatible m would be justified by observations of gold spheres.

But we also want the non-modal statements *not* to be underdetermined. If r is not underdetermined, then r' , the non-modal counterpart of m' , must not be a legitimate competitor of r . This can be the case either if r' is not as well-justified as r , or if it is not actually incompatible with r . But there can be no reason that m' is justified

⁹Perhaps the models of cosmology or climatology do not if they refer to a particular time, the origin of the universe, or to a particular object, the Earth’s atmosphere, but they could inherit their modal force from other theories. This requires justifying adequacy at the theory-level, which will have to wait for the next chapter.

by induction while the weaker r' is not, so this must be because r and r' are not actually in conflict: both can be true at the same time.

How is this possible? Well, in our example, this is possible if there are no gold spheres of more than 1 km in diameter in the whole actual universe. Then it is true that all actual gold spheres are stable (r), and it is also true that all are unstable (r'), because there are none. The universal generalisations are not in conflict, but the corresponding modal statements are, because we are not in a position to know whether such a gold sphere *would* be stable or unstable if it existed. Translated into talk of models, the two models are incompatible, since there are possible situations and contexts (big gold spheres) where the states permitted by the two models (being stable or not) are mutually exclusive, but these situations are not actual. So, the two models M and M' are both empirically adequate in the non-modal sense, but only one of them can be modally adequate, and we do not know which: we have a case of underdetermination. Therefore, it seems perfectly possible that two modal statements should be underdetermined while their non-modal counterparts are not.

5.3.4 Modal Statements are Not Underdetermined

So far, modal scepticism seems vindicated: maybe metaphysical parsimony could incite us to reject induction towards necessity, in particular if it were always possible to find competitors to modal statements that are not in competition with their non-modal counterparts (which remains to be shown). This could persuade us to adopt a non-modal version of empirical adequacy. But let us examine this case more closely: how could we possibly *know* that no big gold spheres exist in the whole universe? What if an alien civilisation in a remote galaxy is creating big gold spheres? What if human scientists in the distant future attempt to create big gold spheres, just to know whether they are stable? What if they form naturally in some parts of the universe? Are there any reasons to exclude these possibilities?

Call r_1 the statement that no big gold spheres exist in the universe, past present and future. Two cases should be distinguished. Either we know r_1 by induction, or we know it by direct observation. The latter case is implausible: r_1 is itself a universal generalisation (all situations are such that there are no gold spheres), and there is no way we could observe all parcels in the universe to make sure that no gold spheres exist. So, we must have acquired this knowledge by induction. But now consider m_1 , the modal equivalent of r_1 , that is, the claim that big gold spheres are *impossible*. If induction justifies r_1 , why could it not justify m_1 ? And if it does, then we have no underdetermination of the modal statements after all, because m and m' are actually compatible: all *possible* big gold spheres are stable, and all are also unstable, because big gold spheres cannot possibly exist in our universe.

What could compel us to assume that r_1 is justified, but not m_1 , is if there could be an underdetermination affecting m_1 but not r_1 . This means that there must be a modal competitor to m_1 , call it m'_1 , such that the corresponding universal

generalisation r'_1 is not a legitimate competitor of r_1 . But we can apply the same argument recursively, and an infinite regress looms: we must assume that a statement m_2 is underdetermined while a statement r_2 is not. Where will it stop?

There are good reasons to think that it will stop at a point where neither m_i nor r_i can be known to be true. In our example, I would argue that we reach this point as early as during the first step: we should not assume that there are no big gold spheres in the universe. But even if the regress does not stop there, it could not go very far.

Bear in mind that induction requires extending a regularity in a sample of situations of a given type to the set of all situations of the same type. The relevant type for r/m is, in our example, situations with gold spheres with a diameter of more than 1 km. r_1/m_1 claim that there is no situation of this type. Therefore, the relevant type for r_1/m_1 must be a less specific type, a “super-type”: it could be situations corresponding to gold spheres of *any size*, the claim being that they are never bigger than 1 km in diameter, and the competing claim being that in some physical configurations, big gold spheres can form. Conditions of application of the corresponding model are less specific, and the domain of application is therefore larger. If we lose specificity at each step in the recursion, at some point, we will reach a completely unspecified type of situation, perhaps something like “any chunk of matter”. We will have a very general model that applies to any situation in the universe.

However, the less specific the type of situation, the more varied its exemplars, and the less representative our sample can be. At some point, induction cannot be vindicated at all.¹⁰ Given the size of the universe, it is very likely that many configurations of chunks of matter that we never observe in our surroundings actually form somewhere, so not observing them does not imply that they do not exist in the universe. This is a reason to assume that there could be big gold spheres in the universe, even if we never observe them.

So, the regression must stop at a point where neither m_i nor r_i are known to be true, which means that both m_{i-1} and r_{i-1} are underdetermined, and neither is known to be true, which means, by recursion, that both m and r are underdetermined. In other words, if a modal statement is underdetermined, then its non-modal counterpart must be underdetermined as well. The reason is the following: if there can be a possible situation where two necessary statements are in conflict, then for all we know, this possible situation might as well be realised, hence the two corresponding universal generalisations are in conflict as well.

My argument rests on an example for the sake of presentation. However, it can be generalised to any universal generalisation, in so far as we consider perfectly general statements, that is, statements that are not restricted to particular objects, places or times in the universe.

¹⁰This can be shown in a Bayesian framework, assuming a principle of indifference to fix prior probabilities for the conceivable states of a given type of situation.

What is relevant for our purpose is a formalisation in terms of empirical adequacy, since our objective is to show that modal empirical adequacy is no more underdetermined than its non-modal counterpart (see Ruyant (2019) for a more general formulation). The argument can be formalised as follows, by considering two incompatible models M and M' (their “permitted states” are mutually exclusive in at least one conceivable context):

1. Hypothesis: we are justified in believing, on the basis of their incompatibility, that at least one of M and M' is not modally adequate. However, we do not know which (they are underdetermined).
2. Hypothesis: we are justified in believing, by induction, that M and M' are extensionally adequate (M is not a competitor of M' for actual situations).
3. If we know, on the basis of their incompatibility, that at least one of M and M' is not modally adequate, then we must know that the type of situation in which they are incompatible has possible instances.
4. If M and M' are extensionally adequate, then the type of situation in which they are incompatible has no actual instance.
5. By 1, 2, 3 and 4, we are justified in believing (by induction and inference) that the type of situation in which M and M' are incompatible has possible instances, but no actual instance.
6. We cannot be justified in believing, by induction and inference, that some type of situation has possible instances, but no actual instance.
7. Contradiction (1, 2, 5, 6): the conjunction of hypotheses 1 and 2 is absurd.

This argument rests on a few assumptions. Assumptions 3 and 4 should be uncontroversial. Assumption 3 presumes that our justification for the fact that at least one of the two models is not modally adequate is their incompatibility in a possible situation. This seems plausible if we do not know which of the two is not adequate, and this is the case that interests us anyway. Step 5 rests on the principle of transmission of epistemic justification by inference, and although there are counterexamples to this principle (Wright 1986), arguably, the present case is unproblematic. The crucial assumption is 6. Its justification is roughly the following: the claim that some type of situation T has no actual instance can be translated into the claim that a model M_1 applying to a less specific type of situation is extensionally adequate. But the adequacy of M_1 is justified by induction only if it has no known competitors, and by recursively applying the same argument, we can show that the modal adequacy of M_1 is just as justified, by the same standard, which would entail that T has no possible instance. The main idea is that every reason we could have to believe that a type of situation is not realised in the universe would be as much reason to believe that it *cannot* be realised.

A direct consequence of the conclusion is the following: if we should believe in the extensional adequacy of a model because it has no known competitor (hyp. 2 is true), then it has no known *modal* competitor (hyp. 1 is false), and we should also believe in its modal adequacy, by the same standard. So, if we have enough reasons to believe that a model is empirically adequate in all actual situations, then we have

enough reasons to believe that it is empirically adequate in all possible situations. We have no reason not to endorse a modal version of empirical adequacy.

5.3.5 *Modal Conflicts in Scientific Practice*

The argument that has just been outlined is rather abstract and formal, and one could perhaps wonder if it is really relevant to real cases. I am convinced that it is.

It is important to note that this argument has a normative aspect. It is an argument about what it is rational to believe. The point is not about whether statements of necessity can be absolutely underdetermined by all actual facts in the universe. The answer is: of course this can happen, for example if, actually, there are no big gold spheres in the universe. The point is about whether it can be rational, at some point in any inquiry, to assume that a statement of situated necessity is underdetermined while the corresponding universal generalisation is not. And the answer is no.

Despite the rather formal nature of the argument, I am convinced that it is relevant for practical inquiry, and for scientific practice in particular. This relevance is strongly connected to the idea that there can be crucial experiments for deciding between competing hypotheses, if not at the abstract level of theories, at least at the level of competing models for a particular type of phenomenon, and that attempting to select competing hypotheses by means of such experiments is part of scientific rationality. This could be illustrated by real-world examples. I will give a few real-world examples later. However, because of the normative nature of the argument, I think that it will be more appropriate to first abstract away from the messiness of scientific practice, and present my case by means of a hypothetical scenario.

For the purpose of illustration, imagine that two groups of scientists each believe in one of our hypotheses about gold and metal spheres conveyed by models M and M' respectively: one group, specialised in gold, has performed many tests on gold and has come to believe that all gold spheres are stable, while the other, specialised in metal, has tested many different metals and believes that since gold is a metal, some of the gold spheres (those that are big enough) are unstable. Would it be rational, in these circumstances, to assume that M and M' are both adequate because they are both justified by different groups of scientists? Would it be rational to assume, as a *direct consequence* of the conjunction of these assumptions, that there are no big gold spheres in the whole universe? This looks absurd, since neither community has ever attempted to justify that claim independently. Would it be rational to conclude that it is impossible to know whether *possible* big gold spheres are stable or not, because we have no “modal telescope” anyway? Should we stop our inquiry there? It seems not.

My intuition is that we could have been justified in accepting either M or M' in isolation. But for the first group of scientists, the fact that the adequacy of M' is justified by another group constitutes new evidence that M might not be adequate after all. It acts as a potential defeater of M , and the inverse is true for the second group. The fact that they are *modal* competitors is enough to put them in conflict.

What is rational, in consequence, is to pursue our inquiry: someone should make the relevant tests to know which of M or M' is adequate. We should try to create big gold spheres.¹¹

What if we discover during the process that gold spheres of 1 km in diameter cannot be created? Then the underdetermination is resolved: big gold spheres are both stable and unstable because they cannot exist. And this resolves the modal underdetermination as well: big gold spheres are impossible in this universe.

What this example purports to show is that in practice, in science or elsewhere, modal incompatibility between general hypotheses is enough to create a conflict to be resolved. So, making sure that a hypothesis is true, or that a model is empirically adequate, involves eliminating all its known modal competitors, which, incidentally, resolves any modal underdetermination. And as a consequence, we have all reasons to accept that all inductively justified knowledge, including the adequacy of a model, has a modal force.

Now, there might be some cases where we have independent inductive reasons to believe that particular types of situations are never instantiated in the whole universe. Arguably, this would be a case where a mere possibility is underdetermined, while the corresponding universal generality can be known. Imagine, for example, that there is not enough gold in the whole universe to ever form any gold sphere of more than 1 km in diameter. Is it enough to consider such cases in order to be a modal sceptic?

A first problem with this strategy is that it could presuppose modal facts. In our example, we could know that there is not enough gold in the universe because we know something about the initial conditions of the universe, and we accept a principle of conservation of matter. But the idea that the initial conditions of the universe, plus some law, limit the kinds of situations that can be instantiated later is modally loaded. Arguably, grounding modal scepticism on modal knowledge is a bit problematic. . . . An answer to this objection could be that the law of conservation of matter is a simple regularity. However, this would require showing that the situations where this law is not respected are possible, but never actually instantiated.

A second problem is that this kind of case is rather limited in scope. It does not warrant modal scepticism *in general*, but only for specific cases, where we have information about some type of situation that is never actually instantiated.

And a third, fatal problem in our context is that these cases only exclude knowledge about possibilities that *could not even happen in our universe*. This means that in the framework of situated possibilities, they should be considered impossible. Remember that we are concerned with alternative ways actual, bounded situations could be, given some fixed background conditions. Arguably, the initial conditions of the universe are part of these background conditions. The alternatives to a situation are situations of the same type: they often contain the same objects, but

¹¹ We could also, of course, remain agnostic. What is irrational is to assume that both models are adequate, and that therefore, pursuing the inquiry is impossible. Modal conflicts are enough to defeat the belief that some general hypotheses are true.

these are configured differently. According to this definition, if there is not enough gold in the universe, then gold spheres of more than 1 km in diameter are *impossible*, and whether such gold spheres would be stable or not should not bother us: the notion of possibility involved here is too metaphysical to ever be tested (see also Sect. 4.2.5).

But for all the more mundane hypotheses, that is, the ones that concern objects to which we do have epistemic access, there is no real problem. If a general hypothesis has no known competitor, then it is legitimate to extend it to all possibilities. And if it turns out that it does have a modal competitor,¹² then we should doubt this hypothesis and make relevant tests. In all cases, the relation of necessity is as well-justified as the universal generalisation.

All these remarks vindicate modal empiricism as the best way of making sense of scientific practice as a rational activity. As we can see, an important aspect of this narrative is that in general, it is possible to decide between competing hypotheses by means of crucial experiments, and that we *should* do this if we want to advance scientific knowledge. The illustration proposed was hypothetical, and the idea that there can be crucial experiments for selecting between competing theories has been notoriously criticised by Kuhn (1962). There is, indeed, a problem of underdetermination by experience that affects this idea, since theories are applied by means of models that incorporate auxiliary hypotheses, and by the mediation of experimental norms that also incorporate auxiliary hypotheses. Kuhn is particularly concerned with paradigm choice in “revolutionary science”. Even accepting this, the idea that competing models within a theory can be selected during periods of “normal science” by means of crucial experiments (accepting Kuhn’s distinction between normal and revolutionary science) should be less controversial.¹³

Roll-Hansen (1989) gives the example of competing hypotheses concerning whether natural selection acts continuously or discontinuously. Experiments on homogeneous lineages confirmed that natural selection acts discontinuously. We have a clear example where two hypotheses make different predictions for a possible type of situation, and Wilhelm Johannsen decided between them by carefully selecting peas so as to implement this possible situation. Another example of crucial experiment that concerns the molecular structure of DNA is given by Weber (2009). Yet another example is the confirmation of the Higgs mechanism, which was actually in competition with other models (Mättig and Stöltzner 2019). In all these cases, there were good (presumably inductive) reasons to prefer one or the other hypothesis, and “unnatural” experimental situations were implemented in order to decide between them. All these cases involve methodological complications, and a lot more could be said about what makes an experiment decisive, but these

¹²I presume that these modal competitors could first be discovered by inference to the best explanation, and then justified by induction. I will argue in the next chapter that inference to the best explanation plays no justificatory role, but at best a strategic role in scientific practice.

¹³These experiments are “crucial” at least relative to background assumptions, assuming an agreement on experimental norms of operationalisation, for instance.

experiments nevertheless convinced the scientific community of the superiority of one hypothesis. We can see that the approach presented in this chapter matches scientific practice in this respect.

5.4 Laws of Nature

Let me finish this chapter with a quick note about the epistemology of laws of nature.

There is an ongoing debate about the laws of nature between Humeans, who assume that these laws supervene on the mosaic of actual facts, and non-Humeans, who assume that they have a “modal force”, to which we could add anti-realists, who assume that the laws of nature do not exist.

One could say that the arguments put forward in this chapter are in line with the views of Humeans and anti-realists when it comes to scepticism about inference to the best explanation. Law-like generalisations are merely the kind of statements that can be justified by induction; no other mode of inference is needed. But my arguments fit with non-Humeanism when it comes to the idea that law-like generalisations have a modal force. It should be clear that there is no contradiction here: law-like generalisations are precisely the ones that can be justified by induction, and this means that they have modal force, because the corresponding laws of (situated) necessity are as well-justified.

But one could worry that the present account cannot really distinguish law-like statements from accidental generalisations, if both are equally well-justified by induction. Accounting for this distinction is, after all, one of the main motives of the debate on laws of nature. The worry could be that accidental generalisations would count as necessary for a modal empiricist, which seems absurd.

5.4.1 *Induction and Projectibility*

This distinction between law-like statements and accidental generalisations is often presented by means of examples, by appealing to our intuitions: intuitively, “no raven moves faster than 30 m/s” is not a law of nature, even if true, but “no object moves faster than 300,000 km/s” is a law. However, both statements have apparently the same linguistic form (Carroll 2008, pp. 2–3). Or the fact that all coins in my pocket are silver coins is accidental, but the fact that all pieces of copper conduct electricity is a law.

Goodman (1954) remarks that there is a strong connection between this distinction and induction. The statement that all coins in my pocket are silver coins cannot be confirmed inductively. It is somehow confirmed every time I observe a new silver coin in my pocket, but this is only because this observation eliminates some possibilities (at least the coin I have just observed is made out of silver). However, observing a new silver coin does not change the credibility that the next

coin that I will observe will be a silver coin: my observation cannot be *projected* on future observations. The same goes when a tossed coin lands on tails three times in a row: this does not affect the probability of the next toss resulting in tails. There is still a 50% chance that it will land on heads. On the contrary, the fact that a piece of copper conducts electricity increases the credibility of the belief that the next piece of copper that I will observe will conduct electricity as well. Being made out of silver is not projectible for the coins in my pocket, but conducting electricity is projectible for copper pieces.

As Hume noted, inductive reasoning rests on a postulate of the uniformity of nature. The lesson of Goodman's remarks is that we are not willing to attribute uniformity to any kind of properties. So, what is the difference between the conductivity of copper pieces and that of silver coins? It could seem that the former is a good candidate for being a law, or a matter of necessity, but not the latter. This would mean that a statement must be assumed to be possibly law-like or necessary *prior* to inductive inferences, because induction is only warranted for law-like or necessary candidates. This is the conclusion drawn by Lange (2000, ch. 4) and Dretske (1977). How do we decide whether a statement is possibly law-like? Perhaps by some inference to the best explanation (which is what Dretske argues): the conductivity of copper pieces could be explained by the nature of copper, while there does not seem to be any good explanation for the fact that all coins in my pocket are silver coin, even if true. If there were an explanation, a reason why non-silver coins were falling outside of my pocket for instance, then induction could be warranted in this case too.

This supports the rationale presented in Sect. 5.3.2: a sequence of observations (conductive copper pieces) would confirm a relation of necessity or law (copper is conductive) by *inference to the best explanation*, and *not* by induction, from which one can *deduce* that all copper pieces are conductive. This would undermine the results of the previous section.

I remain unconvinced by this kind of argument. I would say that they put the cart before the horse. Let us see why.

A useful framework for analysing inductive inference is the Bayesian framework, which consists in the application of Bayes' theorem to derive conditional probabilities between prior and posterior assumptions.¹⁴ In a Bayesian framework, the differences between the inference we make in the case of silver coins, which

¹⁴Bayes' theorem derives directly from probability theory, and it is independent from one's interpretation of probabilities. It requires that probabilities be assigned to hypotheses, but they can be interpreted as subjective degrees of credence or as objective degrees of confirmation for instance, and the probabilities assigned to events, as opposed to hypotheses, can be interpreted differently (assuming an account of their connection with hypotheses probabilities). My use of this framework should not be interpreted as a commitment to any version of Bayesianism, such as subjective Bayesianism, the latter being associated with specific interpretations of probabilities in terms of degrees of credence. I am personally not convinced that subjective degrees of credence play any significant role in science (if they exist at all). However, this framework remains a useful tool for philosophical analysis of rational inferences.

is thought to be an accidental generalisation, and the one we make in the case of copper pieces, a law-like statement, can be addressed in terms of prior probabilities. Imagine we observe ten coin tosses in a row. How shall we assign prior probabilities to possible sequences of observations?

One way is to give equal weight to all possible sequences: ten heads is a priori as likely as nine heads followed by tails, which is as likely as eight heads followed by tails followed by heads, and so on. Since any sequence is equally probable, the first observations tell us nothing about how the sequence will unfold: past observations are not projectible onto future ones. This is what happens in the case of silver coins in my pocket.

Another way to assign prior probabilities is to give equal weight to all property distributions. This means considering that having one heads in total (which corresponds to ten possible sequences) is a priori as likely as having two heads in total (forty five possible sequences), which is as likely as having ten heads (only one possible sequence). In this context, every observation of a heads result affects the probability that future observed results will be heads as well, because it confirms the distributions where the proportion of heads is higher. For example, after having observed nine heads, an application of Bayes's rule gives ten chances out of eleven that the right distribution is ten heads in total, and only one chance out of eleven that it is nine heads in total, which means ten chances out of eleven that the final toss will be heads. This looks like what happens in the case of the copper pieces. As we can see, in a Bayesian framework, Hume's postulate of the uniformity of nature is hidden in a focus on distributions rather than sequences for assigning prior probabilities.

What Dretske's analysis suggests, when transposed to this Bayesian framework, is that opting for an equal weighting of distributions must be justified by a prior inference to the best explanation. An equal weighting of sequences would be the default approach, so to speak, but in some cases, we could be drawn into assigning equal weights to distributions instead, because these distributions could be explained. The problem with this strategy is that it seems to rest on a mysterious intuition concerning which predicates are projectible, or which statements have potential explanations.

In any case, I believe that this is wrong, and that assigning equal weights to universal distributions is the way to go in *all circumstances*: this is not only the default approach, but the only valid one. Assigning equal weights to sequences instead of distributions, as we do when throwing dice or tossing coins, is only the special case we arrive at for particular types of distributions: the ones where all results are equally frequent. We already know, prior to tossing a coin, that the right distribution of results for coin tosses in general is half tails and half heads, and we know this because we (or others) have already performed an *induction* on coin tosses in general. We can *deduce* from this background knowledge that any random sequence corresponding to a sub-set of the larger and evenly distributed set of all coin tosses is as likely as any other (assuming that the larger set is much larger, perhaps infinite). However, nowhere did we need an inference to the best explanation. After a hundred heads, it would be legitimate to suspect that

this particular coin is loaded, so this background knowledge is defeasible. More generally, any enduring pattern (having only silver coins in my pocket for years and years) could prompt the need for an explanation and warrant inductive inference, even before this explanation is found, so it cannot be true that inference to the best explanation comes first.

In sum, there is no need for an inference to the best explanation to select candidates for induction: induction always comes first, and it is always applicable. Only the fact that we rely heavily on background knowledge in daily life could make us think otherwise. The fact that some regularities are accidental is itself discovered by induction. As for explanations, I would argue that they have to do with the unification of various inductively confirmed statements, but this is another topic (see Sect. 3.3.6).

5.4.2 *Laws and Accidents*

Going back to the problem of differentiating statements such as “no raven moves faster than 30 m/s” and “no object moves faster than 300,000 km/s”, if induction is generalised in the way suggested above, it could seem that we lack resources to do so, and that accidental generalisations would have a modal force in the present framework.

Note, however, that the kind of necessity considered in this work is weaker than nomological necessity, because it is relative to our epistemic position and to background conditions (and the exact conditions to which these relations of necessity are relative are presumably unknown). This means, as noted earlier, that the kind of necessity at stake is more permissive than nomological necessity: some regularities might count as necessary in the framework of possible situations, even though they would count as contingent from the perspective of a metaphysics of laws of nature. As argued in Sect. 5.2.2, if, for some contingent reasons, our observations were limited to the surface of the Earth, the idea that all objects in free fall accelerate towards the centre of the Earth at 9.8 m/s^2 would count as necessary (and arguably, assuming fallibilism, it could have been considered as such some time ago), although, as we well know, it is not a universal law of nature. There are many inductive inferences that concern contingent states of affairs rather than laws of nature. They still support counterfactual reasoning. Lange (2000, ch. 4) gives the following example: any person of entirely Native-American heritage is blood type O or blood type A, because all Native Americans are descended from a small, isolated population, and as it happens, allele B was not present in that population. The statement “all Native Americans have blood type O or blood type A” is not a law of nature, because it results from a coincidence, but it can be confirmed by induction, and it can be used for counterfactual reasoning. It can be considered a statement of necessity in the present framework.

This relative permissiveness can defuse the problem of demarcating accidental generalisations from law-like generalisations. I assume that our capacity to imagine legitimate modal competitors for the various hypotheses we could form, and our ability to test them by implementing the relevant experimental situations, is enough to eliminate accidental generalisations. So, I might invent a contrived predicate by which I show, inductively, that all coins in my pocket right now are silver coins. With sufficient imagination, I might propose a general model that only applies to my pocket, without referring to particular places and times. It would look as if I have inductively justified an accidental generalisation to be a matter of necessity. I would argue that such a statement does indeed count as a candidate law-like statement: it is only a very bad one. I can as easily justify the competing statement that all coins in pockets in general are mixed, and this statement has far more occurrences to make its case. My first claim, with its tiny sample, appears to be quite weak in comparison. And if we insisted in considering that both statements are equally justified, we would have a case of underdetermination, but a simple experiment could settle this case: simply add a non-silver coin to my pocket and see if it turns silver.

The idea is that most accidental generalisations could be eliminated, simply because they would not survive the fierce competition to which a good inquiry would submit them. Law-like statements are just the generalisations that survive in this competition. However, in virtue of the permissiveness mentioned above, some cases given in the literature as cases of accidental generalisations could count as necessary connections in the present framework: for example, it might well be *impossible* for ravens to fly at more than 30 m/s, given some environmental constraints. In sum, relaxing the metaphysical conditions for what counts as a law-like generalisation, but maintaining rather stringent epistemic criteria for what is inductively justified, can help produce a meaningful distinction between accidental and law-like generalisation.

The question remains why this statement about ravens, if it is law-like, does not derive from or appear in any scientific theory, as opposed to the claim, of the same form, that no object moves faster than light. My answer would be the following: what induction towards necessity can justify is not fundamental scientific laws directly, but rather observational laws. It is not the empirical adequacy of a theory, but that of a theoretical model. As noted by Cartwright (1983) and Giere (1999), the fundamental laws of scientific theories are rarely applied directly to the world (see Sect. 2.3.1). I consider that they correspond to meta-norms constraining the set of relevant models (Sect. 3.4.1). This means that the empirical adequacy of a theory cannot be justified by a simple induction of the kind analysed in this chapter. This appears clearly in the definition of empirical adequacy proposed in the previous chapter (Sect. 4.3.3): this definition quantifies not only over contexts and situations, but also over models. Something like a meta-induction, an induction on models, will be needed to justify it. This type of induction is presented in the next chapter.

In this respect, the fact that a statement such as “no raven can move at more than 30 m/s” is not a fundamental law, while the statement “no object can move at more than 300,000 km/s” is a fundamental law is not particularly surprising. Note first that the apparent formal similarity between the two statements obscures the fact

that the speed of ravens in question is relative to a particular frame of reference, the Earth. Ravens can certainly move at more than 30 m/s in *some* frames of reference. Now, if we specify “relative to the Earth” in both statements, neither of them is a fundamental law. It is true that the second one is the direct *consequence* of a fundamental law, but granting that the first one is not, the difference between the two statements appears to be extrinsic. If we specify “relative to the Earth” for the first statement, and “in any frame of reference” for the second, we might be able to say that this first statement is not a fundamental law while the second one is. But their forms are now distinct, and it appears that the second statement needs the mediation of a more particular model associated with a particular frame of reference to have any practical consequences for concrete phenomena. The status of this fundamental law is therefore distinct: it acts as an organising principle for a family of models, and its justification will be distinct as well.

I presume that this argument can be applied to other cases: for example, the instability of big uranium spheres, as opposed to the instability of big gold spheres, is a *consequence* of fundamental theoretical laws, and paying attention to these laws would reveal that they have a different status, that of organising principles, because they are not about particular classes of empirical phenomena, but rather they establish relations between theoretical concepts.

In any case, we seem to know a lot of law-like modal facts, about the fragility of glass or the direction of free-fall of falling bodies for example, without necessarily grounding this knowledge in fundamental scientific theories, and this is what matters here: that these facts, with all their modal force, are perfectly knowable from experience.

5.5 Do We Need More?

In this chapter, I have presented the notion of situated possibilities on which modal empiricism is based. Situated possibilities, contrarily to possible worlds, are anchored to actual situations, which brings intensional and extensional limitations. I have also given a few reason to accept their existence: the ubiquity of modal discourse, their role in motivating scientific experimentation and their role in action-based theories of perception.

If one accepts that there are situated possibilities in the world, that is, alternative ways actual situations could be given some natural and environmental constraints, then there is an inductive route towards knowledge of relations of necessity. Realised possibilities can be taken to be representative of merely possible ones, and they inform us about what is or is not possible in this universe. The modality involved is weaker than metaphysical or nomological necessity, but that should not bother us for all practical purposes. Furthermore, there is no reason to *resist* knowledge of situated necessity, because any underdetermination that would affect it, from our perspective, would affect the corresponding universal generalisation just as much: if there is a problematic possibility that contradicts a relation of necessity, then for all

we know, it might well be realised somewhere in the universe. This is in line with scientific practice: a modal conflict between two hypotheses is enough to motivate further inquiry.

Given that the alleged impossibility of knowledge of relations of necessity from experience remains one of the main reasons for modal scepticism, this means that there is no good reason for an empiricist not to embrace natural modalities.

The kind of modality discussed in this chapter is precisely the one involved in the empirical adequacy of theoretical models introduced in the previous chapter. It characterises modal empiricism. So, scientists are justified to believe that their best models are empirically adequate in a modal sense. If only induction is required to justify this adequacy, then there is no reason to think that models correspond in any way to a mind-independent reality, for example, that they correctly describe the essence of natural kinds. This idea would be warranted if their empirical success was in need of an explanation. However, an inductive justification removes the need for explanations: at most what is needed is something like a certain uniformity of nature, now extended to the realm of situated possibilities.

However, this rationale applies to the empirical adequacy of theoretical models, and more should be said to justify the adequacy of theories. The status of law-like statements that can be justified by induction and the status of fundamental laws of theories are distinct: the latter are not about particular classes of phenomena, but rather act as organising principles for the family of models constituting a theory. Maybe the fact that families of models organised by laws and principles are generally empirically adequate is still miraculous from an empiricist perspective, and maybe scientific realism provides the only explanation for this miracle.

I am convinced that modal empiricism gives us enough resources to resist this conclusion, by invoking a different form of induction: an induction on models. This is the topic of the next chapter.

References

- Armstrong, D. (1983). *What is a Law of Nature?* (Vol. 96). Cambridge: Cambridge University Press.
- Barwise, J., & Perry, J. (1983). *Situations and attitudes* (Vol. 78). Cambridge: MIT Press.
- Briscoe, R., & Grush, R. (2015). Action-based theories of perception. In E. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Metaphysics Research Lab, Stanford University, Stanford.
- Campbell, S. (2001). Fixing a hole in the ground of induction. *Australasian Journal of Philosophy*, 79(4), 553–563.
- Carroll, J. (2008). *Laws of nature*. Cambridge: Cambridge University Press.
- Cartwright, N. (1983). *How the laws of physics Lie* (Vol. 34). Oxford: Oxford University Press.
- Dretske, F. I. (1977). Laws of nature. *Philosophy of Science*, 44(2), 248–268. <https://doi.org/10.1086/288741>
- Duhem, P. (1906). *La théorie physique: son objet, et sa structure*. Paris: Chevalier & Rivière.
- Foster, J. (1982). Induction, explanation and natural necessity. *Proceedings of the Aristotelian Society*, 83, 87–101.

- French, S. (2014). *The structure of the world: Metaphysics and representation*. Oxford: Oxford University Press.
- Gibson, J. (1986). *The ecological approach to visual perception*. London: Routledge.
- Giere, R. (1999). *Science without laws. Science and its conceptual foundations*. Chicago: University of Chicago Press.
- Goldman, A. I. (1994). Naturalistic epistemology and reliabilism. *Midwest Studies in Philosophy*, 19(1), 301–320.
- Goodman, N. (1954). *Fact, fiction & forecast* (Vol. 25). London: University of London.
- Hale, B. (2003). Knowledge of possibility and of necessity. *Proceedings of the Aristotelian Society* (Hardback), 103(1), 1–20. <https://doi.org/10.1111/j.0066-7372.2003.00061.x>
- Hawke, P. (2011). Van Inwagen's modal skepticism. *Philosophical Studies*, 153(3), 351–364. <https://doi.org/10.1007/s11098-010-9520-5>
- Hommel, B. (2004). Event files: feature binding in and across perception and action. *Trends in Cognitive Sciences*, 8(11), 494–500. <https://doi.org/10.1016/j.tics.2004.08.007>
- Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, 24(5), 849–878.
- Kratzer, A. (2019). Situations in natural language semantics. In E. Zalta (Ed.), *Stanford encyclopedia of philosophy*. Metaphysics Research Lab, Stanford University, Stanford.
- Kripke, S. (1980). *Naming and necessity*. Cambridge: Harvard University Press.
- Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Ladyman, J., & Ross, D. (2007). *Every thing must go: Metaphysics naturalized*. Oxford: Oxford University Press.
- Lange, M. (2000). *Natural laws in scientific practice*. Oxford, New York: Oxford University Press.
- Lewis, D. (1973). *Counterfactuals*. Oxford: Blackwell Publishers.
- McMullin, E. (1985). Galilean idealization. *Studies in History and Philosophy of Science Part A* 16(3), 247–273. [https://doi.org/10.1016/0039-3681\(85\)90003-2](https://doi.org/10.1016/0039-3681(85)90003-2)
- Mättig, P., & Stöltzner, M. (2019). Model choice and crucial tests. On the empirical epistemology of the Higgs discovery. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 65, 73–96. <https://doi.org/10.1016/j.shpsb.2018.09.001>
- Nanay, B. (2012). Action-oriented perception. *European Journal of Philosophy*, 20(3), 430–446.
- Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, 9(2), 129–154. <https://doi.org/10.1080/713752551>
- Psillos, S. (1999). *Scientific realism: How science tracks truth. Philosophical issues in science*. London, New York: Routledge.
- Putnam, H. (1975). The meaning of 'meaning'. *Minnesota Studies in the Philosophy of Science*, 7, 131–193.
- Roca-Royes, S. (2007). Mind-independence and modal empiricism. In *Proceedings of the 4th Latin Meeting in Analytic Philosophy, Carlo Penco* (pp. 117–135).
- Roll-Hansen, N. (1989). The crucial experiment of Wilhelm Johannse. *Biology & Philosophy*, 4(3), 303–329. <https://doi.org/10.1007/BF02426630>
- Ruyant, Q. (2019). The inductive route towards necessity. *Acta Analytica* <https://doi.org/10.1007/s12136-019-00402-3>
- Schütz-Bosbach, S., & Prinz, W. (2007). Perceptual resonance: Action-induced modulation of perception. *Trends in Cognitive Sciences*, 11(8), 349–355. <https://doi.org/10.1016/j.tics.2007.06.005>
- Stove, D. (1986). *The rationality of induction* (Vol. 23). Oxford: Oxford University Press.
- Strohming, M. (2015). Perceptual knowledge of nonactual possibilities. *Philosophical Perspectives*, 29(1), 363–375.
- van Fraassen, B. (1980). *The scientific image*. Oxford: Oxford University Press.
- van Fraassen, B. (1989). *Laws and symmetry* (Vol. 102). Oxford: Oxford University Press.
- Weber, M. (2009). The crux of crucial experiments: Duhem's problems and inference to the best explanation. *The British Journal for the Philosophy of Science*, 60(1), 19–49. <https://doi.org/10.1093/bjps/axn040>

- Williams, D. C. (1963). *The ground of induction*. New York: Russell & Russell.
- Wright, C. (1986). Facts and certainty. *Proceedings of the British Academy*, 71, 429–472.
- Wright, J. (2018). *An epistemic foundation for scientific realism: Defending realism without inference to the best explanation*. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-030-02218-1>

Chapter 6

Scientific Success



Abstract One of the main motivations for scientific realism is that it would explain the “miraculous success” of science, in particular the successful extension of theories to new domains of experience. After recalling the reasons to doubt the validity of the realist strategy, and in particular, the idea that inference to the best explanation is a principle of justification, this chapter shows that modal empiricism presents us with an alternative way of accounting for the successful extension of theories. This alternative consists in completing van Fraassen’s “selectionist” proposal with the notion of an induction on models: the fact that various models of a theory have been successful justifies, by induction, that other models of the same theory will be successful as well. This particular form of induction is discussed, and I argue that it does not imply that justified theories are true.

6.1 The Miraculous Success of Science

In the previous chapters, I have defined the aim of science according to modal empiricism, and I have defused sceptical arguments against the idea that this aim would be achievable in principle. Modal empiricism is more ambitious than traditional versions of empiricism, because it assumes that scientific theories aim at accounting not only for actual phenomena, but also for possible phenomena. However, could we not be even more ambitious? Maybe science, and its undeniable empirical success, deserves a more optimistic approach. Maybe our best scientific theories do not merely account for possible observations, but they also unveil the fundamental nature of reality.

Relativity theory predicted the deviation of light by massive bodies. The wave theory of light predicted that a bright spot should appear at the centre of a circular shadow created by a punctual source. Quantum theoretical models of atoms predict the emission spectrum of various elements with an accuracy that has not been defeated even though the precision of measurements has been increased by several degrees of magnitude. Such empirical successes are impressive. These examples are often taken to support a particular epistemological position: scientific realism,

according to which successful scientific theories are true, or approximately true descriptions of reality.

As Putnam (1975) puts it, “The positive argument for realism is that it is the only philosophy that does not make the success of science a miracle”. Similar arguments or intuitions had been expressed earlier by Smart (1963) and Maxwell (1970). This so-called no-miracle argument is now generally understood as an inference to the best explanation: the empirical success of our scientific theories calls for an explanation, and the best, or only explanation is scientific realism. What explains the empirical successes mentioned above is that the wave theory of light, general relativity or quantum mechanics are true descriptions of reality, or descriptions that are close enough to the truth. They capture important aspects of its fundamental nature. So their success is no miracle.

Here, we have definitively left the axiological debate about the aim of science for the epistemic debate about its achievements. The two are of course related, and we can understand scientific realism as the claim that science not only aims at true descriptions of reality, but that it has also achieved this aim, at least to some extent.

So, what exactly is it for a theory to be true, to correctly describe reality? As noted in Sect. 2.2.2, there is a tension between the view that theories are best conceived of as families of models, and the idea that they are truth-bearers. A family of structures cannot be said to be true or false, so a theory must also be qualified by a statement asserting that its models have a certain relation with reality. This means specifying “what its models are about”, and “what they say”, and the two must be distinct if misrepresentation is possible (see Sect. 2.3.2). To put it differently, this means providing an account of scientific representation, such as the one developed in Chap. 3.

The two-stage account of representation of Chap. 3 is not perfectly suited for a realist for several reasons. One of them is that contexts play a prominent role: relations between models and the world depend on the purposes of agents, and outside of a context, a model only gives norms of representation that are applicable in various contexts. Norms are not descriptions. Another reason is that this framework assumes that the properties of interest specified by contexts are accessible empirically. Their interpretation is given by norms of experimentation (Sect. 3.4.2). By contrast, a scientific realist would say that theoretical terms purport to refer to natural properties that are well-defined independently of their accessibility, and independently of our purposes and representational activities in general. She would say that theoretical models pretend to describe objects that really exist and have the properties attributed to them by these models, even when these objects are not directly observable. Perhaps she would also say that the dynamical relations between properties postulated by theories are candidates for corresponding to the laws of nature, or that the structures of particular models correspond to causal relations between real objects, and again, these laws and causal relations would be what they are independently of our epistemological position. In sum, she would say that our theories potentially correspond to reality.

The two-stage account of scientific representation presented in this book is not fundamentally incompatible with these ideas. However, this account does not make such postulates, because its purpose is to account for representational activities in a minimal and general way, and we cannot presume that anti-realists are unable to represent. For this reason, a scientific realist could accept that the two-stage account correctly describes the surface features of representational activities, but she would probably be willing to complete this account with realist aspects.

It is not my place to provide a full realist account of scientific representation, but here is a sketch of how this could be done. One could introduce a notion of *realist interpretation* that would not map models and contexts, as is the case in our two-stage account, but models and situations of the world directly. This mapping would not be mediated by norms of experimentation: the symbols used in models would typically refer to real objects and properties of the situation they represent. The realist could take on board the indexical component of representation presented in Sect. 3.3.3, and assume that each theoretical model is associated with a relevance function that specifies the set of legitimate realist interpretations of this model (so contexts would still be present, but they would merely specify salient targets of representation). These relevance functions would determine the domains of application of associated models. We could say that a theory is true if all the legitimate realist interpretations of its models in terms of concrete situations are accurate, but now in a realist sense: the model, now that it is realistically interpreted, corresponds to the represented situation. The structure of objects and properties that it describes is actually instantiated in the target system. Finally, the realist could provide an account of the relation between this realist notion of interpretation and actual representational activities, since the latter involve idealisations and a sensitivity to purposes that is not present in realist interpretations.

There might be other approaches than the one sketched here (one of them is to simply reject model-based conceptions of theories), but whatever the solution adopted by the realist, I assume that something like a realist, as opposed to a merely pragmatic, interpretation of theories is involved, and that it has the following feature: this interpretation transcends our epistemic position. A realist interpretation goes beyond a surface account of the way a theory is used in experimental setups. This is one central feature of the semantic component of scientific realism, so it is non-negotiable. This means that truth is strictly stronger than empirical adequacy, even for the modal version of empirical adequacy defined in Chap. 4. I will assume that theoretical truth entails empirical adequacy, but not the other way around (this assumes realism also concerns our operationalisations, see Lyons (2002) for complications regarding this link). So, from a realist perspective, two theories that are equivalent pragmatically speaking, that are used in exactly the same way and have the same practical consequences in all circumstances, can differ in their realist interpretation, and in so far as they are incompatible, only one of them can be true.

I am not sure that these ideas of realist interpretation and correspondence to the world really make sense (this will be discussed in Chap. 8), but let us assume that they do for the sake of the discussion. What reasons do we have to assume that theories, when realistically interpreted in this way, are true?

As said earlier, the main reason that one can find in the literature is that this is what best explains the impressive empirical success of science. According to the realist, the empiricist merely *states* this success, but she does not *explain* it, and this is why the position is lacking. Another way to express this is to say that empirical success provides a justification not only for empirical adequacy, but also for truth, and the inference involved in this justification is an inference to the best explanation.

The argument has intuitive appeal, but its ramifications are quite complex. I will examine them in this chapter. I will then explain how modal empiricism is able to respond to the realist, by defusing the need for an explanation. The main observation is that empirical adequacy already accounts for the empirical successes of theories, so what we really need is a justification of empirical adequacy, and the question is whether this justification implies theoretical truth. I will provide such a justification by means of an induction on theoretical models. Finally, I will examine what this justification implies, so as to answer this question: how far from scientific realism is modal empiricism?

The two-stage account of representation presented in Chap. 3 will be used in the discussion, and incidentally, I hope that this will prove its fruitfulness for philosophical analysis.

6.2 The No-Miracle Argument

Let us first review the main motivation for scientific realism, the so-called “no-miracle argument”. This argument bears on the role of non-empirical virtues in theory choice and on the status of inferences to the best explanation. The best way of introducing these aspects might be to first examine the problem of underdetermination by empirical evidence, which affects scientific realism.

6.2.1 Underdetermination and Non-Empirical Virtues

The problem of underdetermination is the following (for a general presentation, see Stanford 2017). In principle, any theory or hypothesis has empirically equivalent rivals, which are different theories or hypotheses with the exact same observational consequences. This is a possibility assuming that theories are never mere syntheses of empirical regularities, but rather explanations of these regularities by means of theoretical posits: in principle, there can be more than one explanation for the same phenomena. Assume that only the conformity of the observational consequences of theories and hypotheses with our observations can count as evidence in favour of them. Then there seems to be no reason to accept any theory over one of its empirically equivalent rivals: since their observational consequences are exactly the same, either both are confirmed, or both are disconfirmed by any piece of evidence. This is a problem for the realist: this could mean that the theories selected by

scientists are not necessarily the closest to the truth. This could also be a problem for scientific rationality in general. Perhaps psychological and sociological factors are involved in theory selection, and they might affect the evolution of science in a decisive way. Perhaps the content and validity of well-accepted theories is relative to a social context.

This problem rests on two premises: first, that there are always empirically equivalent rivals to any theory, and secondly, that the confirmation of a theory only depends on its observational consequences. Both premises can be challenged.

The strategy that is generally adopted in order to challenge the first premise is to put constraints on theoreticity. After all, in actual scientific practice, finding one theory that can account for a particular set of phenomena is often quite hard, so this premise seems dubious. Surely, there are trivial ways of producing empirically equivalent rivals to a theory (Kukla 1996): for example, take a theory T and consider the theory T' that states that only the empirical consequences of T obtain, or that they only obtain when someone makes an observation, but not otherwise. T and T' have the same observational consequences, by mere stipulation. However, no scientist would consider such constructs as legitimate rivals of the original theories, since they are clearly parasitic on them.

Note, in this respect, that the debate often seems implicitly framed in a statement-view of theories (see Sect. 2.2): theories are taken to be sets of statements, among which some are observational (or observational statements can be deduced from auxiliary hypotheses), and we are asked to consider alternative sets of statements with the same observational consequences. Transposing the debate to a semantic view, it appears more clearly why constructs of the kind mentioned above are not genuine rivals: they do not consist in proposing new models that would account for the same phenomena, but rather in assuming a different kind of relation between the same models and the world (a different theory of representation). For example, we are asked to assume that the models of the theory only correspond to their targets when measurements are made on the target. What is underdetermined is not the theory itself, but our stance towards the theory, or our meta-theoretical (or semantic) assumptions. This is why such proposals can be presented in full generality, without even presenting the content of the original theory. Or at least, as Leplin and Laudan (1993) observe, it would require much more work to flesh out genuine rival theories along these proposals (modelling the fact that we do or do not measure something, for instance).

This point is particularly salient in the two-stage account of scientific representation presented in Chap. 3 (or its realist alternative presented in the introduction). In this framework, we can say that two models or theories are empirically equivalent if, for any context, the licensed interpretations of the two models or theories bring the same constraints on “permitted states” for the target system. Clearly, proposing an empirical equivalent to a theory T is not as straightforward as stating “ T' says that only the observational consequences of T are true”. It requires proposing different general models with different structures that bring the same constraints in all contexts. As argued by Norton (2008), one could suspect that when this process is too easily achieved, it is because the resulting theory is sufficiently similar to the

original one to be considered a mere reformulation of it. Furthermore, modelling the fact that we do or do not make an observation would amount to switching to a different target of representation, so these theories would not be comparable anymore: we would have a theory of measurements instead of a theory about the original target of representation, for instance.

This casts doubt on the idea that empirical equivalents to our theories can be constructed in a systematic, “algorithmic” way, and this in turn casts doubt on the generality of the underdetermination thesis, at least if this thesis is understood as a thesis of philosophy of science specifically. Indeed, due to their meta-theoretical nature, the proposals examined here seem tailored to raise various kinds of metaphysical scepticism that could apply to any kind of statements, including ordinary ones (perhaps the world only exists while I am looking, etc.), rather than a problem for scientific rationality in particular.

Note, however, that even if these trivial, “algorithmic” cases of underdetermination are rejected, there might still be problematic cases of underdetermination that affect particular theories. Examples of these are a reformulation of Newtonian gravitation in terms of deformations of spacetime rather than forces, and a version of Lorentz’s theory as an alternative to special relativity (Earman 1993). Such examples do not boil down to a radical form of metaphysical scepticism anymore. They do not concern the way our representations relate to reality, and they do not generalise to any theory or statement. Rather, they cast doubt on one particular *realist* interpretation of one particular theory.

Indeed, Norton (2008)’s remark mentioned above also applies to these cases: a formulation of Newtonian mechanics in terms of deformations of spacetime can be considered a mere metaphysical reinterpretation of the same theory, since it can easily be shown that the two theories have the same observational consequences. Following our account of scientific representation, I presume that both versions of Newtonian mechanics will have equivalent models which license the same contextual interpretations. So, what is underdetermined is not the pragmatic interpretation of the theory, but its realist interpretation, and what could be rational, in this context, is to refrain from accepting one realist interpretation over another. This corresponds to an empiricist stance.¹

Another way of responding to this problem consists in minimising one’s metaphysical commitments, assuming that various realist interpretations share a common core. A good contender in this respect is ontic structural realism, according to which our theories should be interpreted in structural terms: they describe “real relations”. The strategy just described has been proposed with respect to the problem of particle identity in quantum mechanics, which is affected by a problem of underdetermination between two interpretations. According to French

¹Note that van Fraassen does not use an argument from underdetermination to defend his constructive empiricism as more rational than other positions. He only uses underdetermination to make sense of the position (van Dyck 2007). According to his voluntarist epistemology, it is not irrational to be a realist, but it is not irrational to be an anti-realist either.

(2014, ch. 2), we can only be realist about the common commitments of these two interpretations, which are structural. For this to count as a version of realism, real relations must be qualified in one way or another, and it remains to be shown that all realist interpretations of theories actually share a core commitment to the specific qualification of the structure that a structural realist must assume. Otherwise, structural realism would seem to be just one among the many available candidates. I will return to this kind of problem and examine structural realism at more length in the next chapter.

As we can see, this problem of underdetermination is specific to realist interpretations. However, there is still the possibility, for the realist, to adopt a more ambitious stance by challenging the second premise of the argument from underdetermination: the idea that the confirmation of a theory only depends on its observational consequences. This premise seems to rest on a naive positivist epistemology. Even if they are empirically equivalent, rival theories (or realist interpretations of a theory) might not be *as well* confirmed by the same observations (Laudan and Leplin 1991). It can be argued that matching our observations is neither necessary nor sufficient for confirmation. For example, a hypothesis can be indirectly confirmed by new evidence because it is part of an overarching theory that is confirmed by this new evidence, even though the hypothesis says nothing about this particular piece of evidence. Psillos (1999, p. 164) gives the example of Einstein's account of the Brownian motion, which was widely taken to confirm the atomic theory although it was not among its consequences.

This points to a more general solution to problems of underdetermination, which consists in claiming that non-empirical virtues, such as theoretical unification, simplicity, scope, coherence with background knowledge or fruitfulness, must be part of a good confirmation theory. These are criteria for being a *good* explanation. So, not only should a theory be explanatory, but it should also be able to provide good explanations: explanations that are simple, fruitful, coherent, etc. With all these constraints on theoreticity and confirmation, perhaps the problem of underdetermination vanishes: one could simply choose the best explanation among empirically equivalent rivals. For example, special relativity can be considered simpler than Lorentz's alternative, because it does not assume that the ether exists, and it can be considered more fruitful because it can be used to build general relativity. This would count as a confirmation of the theory (Acuña and Dieks 2014).

In sum, according to the realist, theories and hypotheses (including realist interpretations of theories) can be justified by a particular type of inference: an *inference to the best explanation*. This mode of inference supposedly goes beyond mere induction: rather than simply justifying a universal generalisation, one could infer, for example, the existence of unobservable entities that *explain* our observations. Whereas induction can at best justify the empirical adequacy of a theory, inference to the best explanation can justify its truth. Such inferences would eliminate problems of underdetermination.

Is it enough to vindicate scientific realism? Well, there are at least three problems with this strategy.

The first problem is that cases of metaphysical scepticism, even if not specific to science, might still be sufficient reason to adopt a more parsimonious interpretation of scientific theories. They could give us reasons to be sceptical with regard to the semantic component of scientific realism, and the idea that something like correspondence truth should be assumed. The fact that this kind of underdetermination is not specific to science does nothing to show that it is not relevant to science. I will address these questions in Chap. 8.

The second problem is that even if cases of underdetermination can be successfully resolved by scientists by taking into account non-empirical criteria, there might exist unconceived alternatives to our theories. Being able to know which of two theories is best confirmed does not mean that we know that one of the two is true, because both could be false. The history of science seems to support this idea. According to Stanford:

We have, throughout the history of scientific inquiry and in virtually every scientific field, repeatedly occupied an epistemic position in which we could conceive of only one or a few theories that were well confirmed by the available evidence, while subsequent inquiry would routinely (if not invariably) reveal further, radically distinct alternatives as well confirmed by the previously available evidence as those we were inclined to accept on the strength of that evidence. (Stanford 2010, p. 19)

By induction, we should assume that we are today in such a position, and we should doubt that our theories are close to the truth.² This problem is not unrelated to the pessimistic meta-induction that will be discussed in Chap. 7.

Finally, a third problem, which will be the focus of this chapter, is that even granting that the non-empirical criteria involved in theory choice are rational and objective, there is no reason to think that they are truth-conducive. At most it can eliminate relativist positions that would claim that theory choice is irrational. But there is no reason to think that inference to the best explanation is a valid mode of inference, rather than a heuristic or pragmatic device.

6.2.2 *Against Inference to the Best Explanation*

Arguably, resorting to inference to the best explanation constitutes the main dividing line between scientific realists and empiricists,³ and there are various arguments against the validity of this mode of inference that I will not repeat here (see for example van Fraassen 1989, ch. 6) (Cartwright 1983, ch. 1).

Note that when considering indirect confirmation as a solution to the problem of underdetermination, one particular virtue, unificatory power, plays a central role.

²For critical responses, see Chakravartty (2008); Godfrey-Smith (2008); Devitt (2011).

³However, Ghins (2017) has proposed a different defence of realism about unobservable objects that does not rest on inference to the best explanation, but on causal inferences. A similar strategy is adopted by entity realists, and it will be briefly discussed in Chap. 7.

A typical case of indirect confirmation is when a hypothesis is the consequence of a more general theory. In such a case, any evidence confirming the theory indirectly confirms the hypothesis. As observed by Hempel, this principle cannot be applied to any random “theory”, for example any conjunction of two statements, otherwise any evidence would confirm any hypothesis (Okasha 1997). There are criteria of theoreticity for this principle to be sound. Not any set of statements constitutes a potential scientific theory. In particular, the theory must unify various phenomena into a coherent whole.

So much is perfectly acceptable for an empiricist, because indirect empirical confirmation is *still* empirical. It can indeed be shown that a certain notion of unity “boosts” confirmation in Bayesian inference (Myrvold 2003). For example, there is a systematic concordance between the retrograde motion of planets and their positions relative to the sun. In the Ptolemaic model of the solar system, this concordance is a mere coincidence: epicycles that do not respect it are a priori possible. However, this concordance is a direct consequence of the Copernican model. We would naturally say that the Copernican model *explains* the fact that retrograde motion occurs when a planet is in opposition to the sun. Myrvold shows that assuming equal prior probabilities for the two theories, the Copernican model ends up having higher posterior probability given empirical evidence, simply because it unifies these phenomena in a precise, probabilistic sense (their correlation is higher assuming the theory). In daily life too, we are often led to give more credence to hypotheses that explain away coincidences.

This means that the unificatory power of a theory can be considered an empirical criterion, in so far as it can be confirmed by observations. Remember that in the definition of empirical adequacy introduced in Chap. 4, a criterion of unification is implicitly involved in the notion of relevance. Such a criterion is enough to eliminate some cases of underdetermination, and it was introduced precisely for this purpose in the context of ad-hoc hypotheses. Ad-hocness and theoretical unification are related (one could say that the concordance between retrograde motion and the position of the sun is ad-hoc in the Ptolemaic model). So, the modal empiricist accepts that unificatory power is an important virtue, and this virtue is indeed part of what makes a model or a theory empirically adequate.

However, unificatory power, understood in this probabilistic sense, and in so far as it only concerns our observations, is not enough to infer that a theory is true. The realist interpretation of a theory is still underdetermined assuming this criterion, because two realist interpretations of the same theory are exactly as unificatory in this sense (Myrvold (2016) is clear that his notion of unification does not support an inference to the best explanation). This is where the scepticism that characterises empiricism starts to bite. When it comes to metaphysics, for example the existence of unobservable entities or of natural kinds, there is no reason to favour one interpretation over another, as long as they are empirically adequate. It does not matter whether Newtonian mechanics is interpreted in terms of deformations of spacetime or in terms of forces of gravitation. It does not matter whether the ether exists, or whether quantum systems are really point particles, multi-fields or collections of events (Belot 2012). For the empiricist, inflationary metaphysics starts

exactly where empirical adequacy, taken to include unificatory power, becomes ineffective as a criterion of choice between competing theories.

An empiricist does not necessarily deny that non-empirical criteria are involved in theory choice. But she will generally consider that they are pragmatic or strategic rather than justificatory. Take simplicity, for instance: a simple theory could be easier to handle, or a good strategy for achieving empirical adequacy rapidly could consist in testing simple hypotheses first before devising more complex ones in case they fail. In a context where observations are prone to errors, models with too many adjustable parameters run the risk of fitting the noise of a particular data sample rather than the signal (this is called overfitting), and for this reason, models with fewer parameters are more likely to make good predictions when used with new data (Forster and Sober 1994). This can be a pragmatic reason to opt for simpler models. But there is no reason to think that nature is simple, and simplicity cannot override empirical adequacy. It can easily be abandoned if required, and the history of science does not look like an ineluctable progress towards simpler theories. There is no reason to think that simplicity or other aesthetic criteria are indicators of truth, even if they result in more satisfying explanations.

It cannot be denied that inferences to the best explanation are used in science, and in everyday life as well. According to a modal empiricist, they can be warranted so long as only unificatory power is involved. In this case, they are merely disguised forms of induction. However, in so far as they would go beyond induction, they should be considered, at best, heuristic devices that can be used to formulate new hypotheses, or strategic devices that can help us decide which hypotheses to consider as a priority. Hypotheses selected by means of non-empirical criteria should then be put to the test in order to be justified by experience: being simple or fruitful or respecting any kind of metaphysical view is no justification of any sort. Empirical tests only confirm empirical adequacy, and when it comes to the realist interpretation of empirically equivalent theories, they become ineffective; only pragmatic considerations remain, and we should not assume that any of these realist interpretations are true.

Interestingly, this view of inference to the best explanation is actually how Peirce, who introduced the concept under the term “abduction”, viewed it (Nyrup 2015). He claims, for example, that “[a]bduction [...] furnishes the reasoner with the problematic theory which induction verifies” (Peirce 1931, 2.776).

Note that this downplaying of the role of inference to the best explanation by no means amounts to denying that theories are explanatory. However, an empiricist will typically deny that explanatory power and truth are connected. After all, as observed by van Fraassen (1980, p. 99) and Bokulich (2016), the standard explanation for tides in oceanography is given by a superseded theory, Newtonian gravitation, and not by its supposedly more veridical successor, relativity theory. The widespread use of idealisations in science seems to indicate that the best explanation for some phenomena is not, in general, the closest to the truth. Arguably, idealisations foster non-empirical explanatory virtues, such as simplicity and cognitive tractability. So, idealised models explain better than their non-idealised counterparts, but this is achieved by deliberately moving away from the truth.

According to Bokulich, falsehood and fictions can be considered explanatory so long as they capture patterns of counterfactual dependence between relevant variables. If these relevant variables are empirically accessible, then the modal empiricist will conclude that theories are explanatory so long as they are empirically adequate. For a modal empiricist, explanatory power is not a criterion for justification, but a mere by-product of empirical adequacy (see Sect. 3.3.6). There is nothing more to scientific explanations than an application of theories in particular contexts, with particular variables of interest (van Fraassen 1980, ch. 5).

There is no doubt that some realist interpretations or formulations of theories have more virtues than others. Some formulations are heuristically fortunate. They can lead scientists to propose new hypotheses or to construct new theories which turn out to be empirically successful. Hamiltonian and Lagrangian formulations of Newtonian mechanics have played this role, for example. McAllister (1989) and Salmon (1990) have suggested that the non-empirical virtues that are constitutive of good explanations could be justified by a meta-induction on the basis of the history of past successes. Non-empirical virtues would be confirmed as virtuous by the fact that the theories and hypotheses displaying them have proved successful in the past.⁴ Even if this is the case, this does not show that these virtues are truth-conducive: the past successes that confirm them, as well as the future successes we expect from taking them into account are all empirical successes, so at most we can say that these virtues are conducive to empirical adequacy. Conformity with experience is always the final judge.

Take relativity theory for instance. According to Zahar (1973), this theory was considered heuristically superior to Lorentz's theory, and Einstein's programme was adopted for this reason by a substantive number of scientists as a research programme. However, as Zahar further explains, this research programme did not really supersede other programmes until the predictions of general relativity turned out to be confirmed by experience. So, non-empirical virtues do play a strategic role, but experience has the final word.

For all these reasons, the realist owes us a positive argument to convince us that non-empirical virtues are indicators of truth, and that inference to the best explanation is a valid form of inference. The realist should tell us why the empiricist approach towards non-empirical virtues presented here would be insufficient. This is where the no-miracle argument comes into play.

⁴Salmon proposes that they affect prior probabilities of hypotheses in Bayesian inferences performed by scientists. As explained in footnote 14 of Chap. 5, I personally doubt that such prior probabilities, interpreted as subjective degrees of credence, play a significant role in science.

6.2.3 *The No-Miracle Argument*

The main realist argument is the no-miracle argument, according to which the success of science would be a miracle if our best theories were not true, or approximately true. The argument has intuitive appeal: who would deny that DNA exists as described by scientists, given the huge amount of confirmation that its existence receives every day? Of course, everything depends on what we mean by “exists”, and an empiricist could well accept that DNA exists in a sense, “for all practical purposes”, without giving too much ontological weight to this statement (this is a semantic issue that I will address in Chap. 8). The scientific realist will want to be more positive, and claim that DNA exists as described by scientists in an absolute sense, as an objective category of nature that does not depend on our representational activities or perspective on the world. However, fleshing out an argument from success in order to justify this positive attitude is more difficult than it seems.

A naive version of the no-miracle argument could run as follows:

1. Theory T is highly empirically successful.
2. The success of T would be a miracle if T was not true.
3. There are no miracles. Therefore, T is true.

But close examination of this argument shows that it is question-begging. Take T to be quantum theory, for example. We could have the realist intuition that it would be a miracle that quantum theory works so well if it were not true, so the truth of T explains its success. However, the explanandum “quantum theory is very successful” can be rephrased as “all the observational consequences of quantum theory are true” (considering ideal empirical success in a statement view of theories for the sake of simplicity). This, in turn, can be rephrased by listing all these empirical consequences, by using the disquotational property of truth (“ p is true” is equivalent to p). The explanans “quantum theory is true” can be reformulated, using the same disquotational property, by presenting the content of the theory. So, what premise 2 states is ultimately something of this kind (taking as an example the explanation of an emission spectrum in quantum physics): it would be a miracle that the spectral lines of atoms of type X have frequencies Y and Z if it was not the case that electrons exist, that they emit photons when changing their energy states with frequency corresponding to the difference of energy between the states, and that only energy states corresponding to frequencies Y and Z are available for atoms of type X . Basically, all the realist does with this naive argument is rehearsing the explanation provided by the theory, and claiming that this is the only possible explanation for the phenomena we observe, and that therefore it is true (for a similar remark, see Levin 1984).

This is a simplification, because in practice, models rather than theories explain and have empirical consequences, and models have a certain autonomy from theories. But acknowledging this makes the realist argument harder to uphold, if anything, because having at least one successful model among the candidate models

of a theory is easier to achieve, and seems less miraculous than if only one model was imposed by the theory. However, even this simplified version fails to provide a convincing argument.

The problem is that this idea that our theories are the only explanations for the phenomena they account for is implausible. The phenomena we observe are not particularly miraculous in themselves, and it would not be a miracle if our explanations of them were actually false: there might be alternative explanations, perhaps unconceived ones, for the same phenomena. Even weakening the argument, and claiming that our theories are not necessarily the only, but at least the best explanations for these phenomena, and are therefore true, would beg the question, because this is precisely what the anti-realist denies. We have seen that idealised models explain better than non-idealised ones, and that superseded theories can provide good explanations. Now some phenomena, such as Fresnel's bright spot, might be genuinely unexpected, and they could seem somehow "miraculous" without a good explanation for them. However, I would say that they are unexpected only in light of past theories or common-sense assumptions. Finding an explanation for a phenomenon, however unexpected, does not make this explanation *ipso facto* true.

This is the reason why realists generally put emphasis on novel predictions. What is remarkable is that Fresnel's bright spot, although unexpected, was predicted by the wave theory of light *prior* to actual observations of the phenomenon. The intuition here is that the wave theory of light must have got something right about the behaviour of light, otherwise, how could it have predicted a phenomenon that was unexpected at the time it was developed?⁵ More generally, what seems miraculous is that a theory that was designed to explain a given type of phenomena can be extended to new contexts and still be successful. This includes an extension to: (i) other types of phenomena that were never observed before, (ii) unprecedented levels of precision, (iii) new operational means of accessing the same theoretical entities (Hacking 1983) (sometimes referred to as corroboration), and (iv) new theoretical developments that will be empirically successful (Boyd 1980; McMullin 1984).

Point (iv) is related to the conjunction argument. Assuming that two theories T_1 and T_2 are true entails that their conjunction is true, but the fact that T_1 and T_2 are empirically adequate does not entail that their conjunction is. So, the empiricist would be unable to explain the fact that combining successful theories or models leads to new empirical successes.

In contrast with the naive version of the no-miracle argument presented above, this approach rightly moves to meta-theoretical considerations that are more relevant for the discussion (this move is due to Boyd). With this approach, we are now focusing not only on the phenomena to be explained or predicted, but also on the methods of selection of theories and hypotheses, presumably inference to the best

⁵However, Hitchcock and Sober (2004) argue that this new prediction did not play a prominent role in the acceptance of the wave theory of light, because the accuracy of the theory for diffraction phenomena was impressive enough.

explanation. The miracle in need of an explanation is not the phenomena themselves, but the extraordinary success of the methods used by scientists for selecting their theories. These methods prove to be ampliative: the theories thus selected can be successfully applied to new domains of experience, and, for example, predict phenomena that were unexpected prior to their observation. Finally, these methods are not merely inductive: new theories do not generalise our observations, but they *explain* them, so inferences to the best explanation are involved.

Focusing on selection methods and novel predictions leads to a more subtle version of the no-miracle argument, which Psillos (1999, ch. 4) calls a “meta-abductive strategy”. It can be formulated as follows.

1. Our best scientific theories are selected by inference to the best explanation on the basis of available evidence.
2. Our best scientific theories can be extended with success to domains of application, levels of precision, etc. beyond that of the evidence on the basis of which they were first selected.
3. The best, or only, explanation for these successful extensions is that inference to the best explanation is truth-conducive.
4. Therefore, inference to the best explanation is truth-conducive, and therefore, theories selected by inference to the best explanation are likely to be true or close to the truth.

What we have here is a justification of inference to the best explanation by inference to the best explanation, hence the “meta-abductive” qualification (whereas the first, naive argument was merely a first-order inference to the best explanation). One problem is that this justification is circular for this reason, but according to Psillos, the circularity is not vicious.

The argument is expressed in terms of theories, but let us note, again, that models rather than theories directly make predictions, so the argument should be qualified: are we talking about selection methods and successes for models or for theories? I presume that the no-miracle argument could apply to both.⁶

In the case of novel predictions, the theory is successful, because novel predictions involve the extension of the theory by means of different models. The no-miracle argument would bear on the selection of what I have called meta-norms of relevance, which correspond to theory-level laws and principles that constrain the construction of models (see Sect. 3.4.1), to the effect that the virtues involved in the selection of these laws and principles (simplicity, etc.) are truth-conducive, since the norms of representation that the selected laws convey often lead to new empirical successes. The same goes for new successful theoretical developments, if these are based on theories, although single models can lead to new successful developments as well.

⁶However, talking about inference to the best explanation could be inaccurate in the case of theory selection if theories are not directly explanatory. Perhaps talking about inference to the best *type* of explanation would be more accurate.

Note that in this case, the argument cannot work at the level of a single theory. Arguably, the fact that a single theory has been successfully extended is not sufficient to infer that it was selected by a truth-conducive method. If the wave theory of light had been revealed to a scientist in a dream, we would probably suspect that its successful prediction of Fresnel's bright spot was just a lucky strike, and we would not necessarily infer that the theory is true. This is because revelation is not systematically reliable. So, what is at stake is not the success of one particular theory, but the reliability of scientific methodology in general. This version of the no-miracle argument does not provide a *theory-level* explanation, but a *science-level* explanation, and the truth of particular theories is not given as an explanation for their success, but deduced from the prior conclusion that inference to the best explanation is truth-conducive in general. This will be important for the arguments of the next section.

In the case of unprecedented levels of precision, on the other hand, the focus is on particular theoretical models that are successful when applied in new contexts, so the argument bears on the virtues involved in the selection of models, and the domain-specific norms they convey (although these are constrained by meta-norms). In this case, it is possible to devise a no-miracle argument at the theory level that explains the successful extension of models to new contexts. Remember that theories can be associated with specific forms of explanations, for example, evolutionary explanations in biology (Sect. 3.4.1). These forms of explanation involve general principles, including what I have called compositional norms, by which complex models are constructed from simpler ones (Sect. 3.3.5). The idea would be the following: the success of *models* when applied to new contexts (for example, new levels of precision) is impressive, and the best explanation is that the inferences to the best explanation from which these models are constructed, which consist in applying the laws and principles of the theory, including rules of model composition, are truth-conducive.

Finally, in the case of various empirical accesses to the same theoretical entities, maybe we could have a look at norms of experimentation (Sect. 3.4.2), assuming an inference to the best explanation is involved as well in the selection of these norms. We can see that the two-stage account of representation, with its various levels of norms, brings interesting distinctions that can help make sense of the different levels at which a version of the no-miracle argument applies.

Is truth-conduciveness of inference to the best explanation the best, or only, explanation for the successful extension of scientific theories to new domains? In this chapter, I will argue that this is not the case. I will argue, more precisely, that the empirical success of science can be entirely justified by a specific form of induction, and that inference to the best explanation need not be invoked. Or rather, if this had to take the form of an alternative explanation (because the debate is expressed in these terms), the explanation would be something like a postulate of uniformity of nature: just enough to justify the kind of induction that is required.

6.3 An Inductivist Response

Let us see how the modal empiricist can respond to the realist demand of accounting for the “miraculous” success of science without resorting to inference to the best explanation. The alternative “explanation” provided by modal empiricism is an extension of the one suggested by van Fraassen, so let us first examine it.

6.3.1 *The Selectionist Explanation*

Van Fraassen (1980, p. 40) is not convinced by the no-miracle argument. He has suggested an alternative “selectionist” explanation:

I claim that the success of current scientific theories is no miracle. [...] For any scientific theory is born into a life of fierce competition [...]. Only the successful theories survive—the ones which in fact latched on to actual regularities in nature.

This (brief) explanation for the empirical success of theories can be interpreted in terms of a selection bias: the success of theories seems miraculous only because we only ever consider the theories that were retained because of their successes. However, it is likely that many theories and hypotheses selected by inference to the best explanation were quickly rejected. So, the truth-conduciveness of inference to the best explanation cannot be the right explanation. The selection method that is successful is not inference to the best explanation, but simply the elimination of non empirically-adequate theories, and the right explanation for the success of our current scientific theories is that they were retained precisely because they were successful. Note that in principle, this explanation could also apply to the successes of models when used in new contexts, with new standards of precision, etc. These successes would be explained by the fact that only the models that pass empirical tests are eventually retained. In this picture, explanatory virtues would play a strategic or pragmatic role rather than a justificatory one.

Van Fraassen’s argument has been criticised for only providing a “phenotypic explanation”, while scientific realism also provides a “genotypic explanation” (Psillos 1999, pp. 93–94). The distinction between these two types of explanation can be illustrated as follows: imagine that during the final game of a tennis tournament, someone asks you to explain why the players are so good. A phenotypic explanation would be: because they were selected by a fierce competition, the tournament itself. Only good players can make it to the final game. A genotypic explanation would be: because these players are fast, muscular, reactive, smart, etc.

The argument is that scientific realism would be a better position than anti-realism because it provides a genotypic explanation for the empirical success of scientific theories: they are successful because they are true. Such explanation is not available to the anti-realist. However, this argument is flawed (and it was actually anticipated in van Fraassen (1980, p. 40, footnote 34)).

In the context of the present discussion, the distinction between a genotypic and a phenotypic explanation corresponds exactly to the difference between the naive, theory-level no-miracle argument presented in the previous section and the more elaborate, science-level no-miracle argument presented later on.

In the genotypic case, we want an explanation for the success of particular theories or models by means of their specific features. For the realist, the relevant explanatory feature is truth. However, being true does not seem to mean that the theories concerned have particular observable features in common (compare to: “all players winning tournaments are muscular”). Rather there is a distinct explanation for each theory’s success, which, as already observed, boils down to claiming that theories explain the phenomena they predict. The empiricist can accept as much: our theories indeed explain the phenomena they predict. So, the empiricist actually has a genotypic explanation for empirical success, and it turns out to be the same as the realist one! The only difference is that the anti-realist denies that constituting an explanation implies being true in the case of scientific theories. Empirical adequacy is enough for scientific explanations. As exemplified by idealised models, a theory or model can be empirically successful and explanatory without being true (see also van Fraassen 1980, ch. 5.1).

Maybe what the realist demands is to provide a *true* genotypic explanation for empirical success. In this case, what the realist actually demands is merely to accept that a good explanation must be true. However, as explained before, this demand is question-begging, because the idea that our best theories must be true is precisely what the empiricist denies. One cannot blame empiricism for not being realism.

As for the more subtle version of the no-miracle argument, it is focused on the way theories are selected, and the explanation for success involves our selection methods. The question that is answered is not “why is this theory in particular so successful?” but rather “why are we in possession of theories that are so successful?”. What is called for is not a genotypic explanation, but a phenotypic one, just like van Fraassen’s explanation. So, van Fraassen’s explanation is a legitimate alternative.

What could give credit to van Fraassen’s explanation is the observation made above that models rather than theories explain and predict. It entails that the successful extension of theories to new domains might not be as miraculous as it seems, because there is a certain flexibility in model construction. If the way of applying theories to new phenomena was entirely determined by the theories themselves, then success would be harder to achieve, and maybe miraculous, but this is not so. Combining the idea that models are selected for their success when extended to new contexts, and then the idea that theories are selected for having successful models when extended to new domains, we might well explain why contemporary theories are so successful.

However, van Fraassen’s explanation might be insufficient in one respect: it does not explain the continuing success of theories after they were first selected, or “why theories that were selected on empirical grounds then went on to more predictive success” (Lipton 2004, p. 195). Relativity theory superseded Newtonian gravitation after successfully predicting the deviation of light by the sun. It could have failed,

and it was retained because it succeeded. There is no miracle here if we balance this success with all the failures that other now forgotten theories have faced in the history of science. However, relativity theory has been extended to new domains since then, and it has never failed to be successful. It was no longer in a “fierce competition”. This, perhaps, is still miraculous and deserves an explanation that the empiricist does not provide. So, one advantage remains for scientific realism.

In order to address this issue, I propose completing van Fraassen’s explanation with an inductive account.

6.3.2 *Induction on Contexts*

What needs to be explained is the continuing empirical success of theories after their initial selection, in particular when it comes to novel predictions. Remember that empirical adequacy is an ampliative notion: it is an expression of ideal success. So, in effect, a justification of the empirical adequacy of a theory (or of a future evolution of the same theory) on the basis of its initial successes is largely enough for our needs. If we are justified in believing that a theory is empirically adequate, that is, accurate for all *possible* applications, then we are also justified in believing that it can make novel predictions, because these are among the possible applications just mentioned. We can see that the focus of the debate on explanations is not entirely fair: any justification is enough to claim that empirical success is no miracle; this justification need not be an inference to the best explanation, and demanding a further explanation for a justified statement is misplaced.

For the realist, the justification of empirical success is indirectly given by the truth-conduciveness of inference to the best explanation, which implies that theories are likely to be true, which implies, in turn, that they are likely to be empirically adequate and to lead to novel predictions. What I propose, as an alternative, is a direct inductive justification of empirical adequacy: the first successes of a theory give us direct inductive reasons to assume that it is empirically adequate, and that it can therefore lead to novel predictions, without requiring a detour by truth.

Let us start by considering theoretical models. Per the definition of Sect. 4.2.6, a model is empirically adequate if its legitimate interpretations are accurate in all situations and contexts. This suggests that the empirical adequacy of a model can be justified by a simple induction on situations and contexts.⁷ Induction consists in extending a particular property from a sample of entities to a population of the same type. Here, the relevant type of entity is the kind of situation and context for which the model has licensed interpretations, for example, hydrogen gases in

⁷For a modal empiricist, the relevant range of situations is the set of all possible situations, so the claim that a model is empirically adequate is a statement of necessity rather than a universal generalisation. I have argued in the previous chapter that the same kind of inductive justification is available for both types of statements, but in any case, the arguments in this section do not rest on this modal aspect.

various conditions where Bohr's model is deemed applicable, or a star orbiting around a black hole represented by a relativistic model, and the property attributed to situations of this type corresponds to the conditions of accuracy of the model, for example, emitting and absorbing rays of light at particular frequencies or having a certain trajectory (see Sect. 5.2.2). Note that these conditions of accuracy are a function of contexts, so the entities we are concerned with (the truth-makers for the property that is predicated) are presumably perspectival. We are interested in some properties of the situations, and not others, sometimes assuming some manipulations. This perspectivity can be more or less innocuous depending on whether active interventions are implied by the properties of interest. However, this does not affect the argument. After having observed that the situations and contexts where the model is applicable are such that the conditions of accuracy of the model are fulfilled, one infers that the model would be accurate in all situations and contexts of this type, which is to say that the model is adequate. Hence, the adequacy of a model can be justified by a simple form of induction.

Is it enough to respond to the no-miracle argument? An adequate model correctly accounts for empirical regularities associated with a given type of situation and context, and it could well be considered as a "black box" for making predictions in these contexts. There is no need to assume that any of its realist interpretations are true. However, as Psillos (1999, p. 73–74) notes:

To be sure, 'black boxes' and the like are constructed so that they systematise known observable regularities. But it does not follow from this that black boxes have the capacity to predict either hitherto unknown regularities or hitherto unforeseen connections between known regularities.

This is certainly a valid observation. Nothing justifies the assumption that a model could predict "hitherto unknown regularities", at least if we restrict ourselves to an induction on situations of a given type. Note, however, that an induction on contexts already seems to provide something more complex than a bare regularity.

Remember that contexts encapsulate properties of interest and degrees of precision. Take for example a model of optics that predicts that light rays are reflected by surfaces at the same angle as the incident angle. After having observed that the model is accurate for incident rays of 32 degrees, and 56 degrees, and so on, it seems that we can infer, by induction on contexts, that it will be accurate whatever the angle of the incident ray. This already looks like a "connection between known regularities", if each angle is associated with one regularity. The same could be said about the model of a star orbiting around a black hole if the mass of the black hole and initial velocity and position of the star are left unspecified, and determined by the context. Furthermore, in the case of our optical model, after having observed that it is accurate for a precision of one degree in our measurements, we could be tempted to infer, by induction, that it would be accurate for higher degrees of precision, too. So, it could seem that an induction on contexts can already help us respond to one version of the no-miracle argument: the one that concerns the successful extension of models to new contexts, and notably to new standards of precisions.

I think that this type of induction on contexts is common in scientific reasoning. One example of this is Galileo who proposes to “observe what happens in the thinnest and least resistant media, comparing this with what happens in others less thin and more resistant” so as to infer what would happen in a medium completely devoid of resistance, to which he had no access (see Sect. 5.2.3). This is an induction on contexts.

However, inductions of this kind seem suspicious, in particular when it comes to degrees of precision. In the case of light reflection, our model postulates a simple linear relation between incident rays and reflected rays, but there is an infinite number of curves that could fit the data available so far, and they are all equally justified by the same induction. Why think that the right curve is a line? Does it not amount to assuming that the right curve is simple? Then should it not be considered an instance of inference to the best explanation?

A partial response to these questions lies in the approach towards induction presented in Sect. 5.3.1. I assume that a statement is justified by induction in so far as no legitimate competitor that is as well confirmed is currently available. Some competitors are illegitimate. I had framed this idea in the context of projectible predicates, assuming that *grue*, *blite* (white before *t* and black after *t*) and the like are not projectible predicates, so that “all swans are blite” does not constitute a legitimate competitor to “all swans are white”. We could propose a similar rationale in the context of curves: maybe not any random curve is a legitimate competitor to a line, because some of them are contrived. In this respect, note that a general model incorporates norms of relevance that should be consistent across uses and avoid ad-hoc postulates, which already eliminates potential curves, such as positing a particular result for only one range of initial conditions. In effect, we are providing a response to a problem of underdetermination that is similar to the ones examined in Sect. 6.2.1.

However, this response is not sufficient, because we could suspect that legitimate competitors remain. For example, a sinusoid approximates a linear relation for small values. Perhaps assuming a sinusoid or a line makes no difference when it comes to interpolating between values if they are small enough, so that an induction on contexts can justify a bit more than simple regularities, but at least it makes a difference when increasing degrees of precision. So, there is no inductive justification that a model will continue to be empirically adequate for higher degrees of precisions. How, then, shall we explain that scientific models continue to be successful?

I would simply respond that they do not always continue to be successful. Very often, linear relations are postulated to account for empirical phenomena, and then second-order corrections are added when the precision of measurements is increased. An inductive justification of empirical adequacy correctly predicts that successful models could well necessitate corrections for new levels of precision. If simplicity is involved, it is merely a strategic or pragmatic criterion of choice, but no particular expectations should be made concerning the extension of our simple curves to new contexts involving higher levels of precision.

The problem with this response is that we can no longer account for the impressive cases of continuing success that are put forth by realists. One example is the prediction of the emission spectrum of various elements mentioned in the introduction of this chapter. Note that these are not cases of curve fitting problems. The mathematical structure of a spectrum is not that simple, and it was motivated by quantum theory (arguably, curve fitting by statistical methods is involved in the construction of data models rather than in the construction of theoretical models). This means that the response just given is still correct: in the case of mere curve fitting, there is no justification for the belief that a model would still be adequate for higher levels of precision, and very often, they are not. However, models constructed on the basis of a theory often continue to be adequate at higher and higher levels of precision. It is not surprising that a simple induction involving a single model cannot do justice to this kind of success if theories are involved, but we still lack an account of it.

The other kind of success that cannot be accounted for by an induction on contexts is the case of novel predictions, because they do not involve the application of the same model in new contexts, but the construction of new models adapted to new types of situations. So, an induction on contexts and situations for a single model will not do. We need something more to do justice to the no-miracle argument.

6.3.3 *Induction on Models*

It seems that simple induction can justify the continuing empirical success of models for the type of phenomena that they were designed to account for, but that it cannot be used to justify much more. This limitation makes this kind of induction unable to address the examples put forth by scientific realists.

The solution that I propose in order to solve this problem is quite simple: let us just push induction one step further. We are looking for a justification of the empirical adequacy of theories. A theory is empirically adequate if all its models are. So, let us proceed to an induction on models. We have a sample of models respecting the norms (laws and principles) of the theory, and they are empirically adequate. By induction, we could expect other models respecting these same norms to be adequate as well. This predicts that the theory will be successfully extended to new domains of application, at least if the sample is large enough (if the theory has already made novel predictions,⁸ for instance).

⁸I agree with Hitchcock and Sober (2004) that successful novel predictions are desirable only in so far as they guarantee that the theory did not overfit the new data. That is, the novelty of predictions is not desirable for its own sake when it comes to justification.

This solution can be combined with van Fraassen's explanation, so as to explain why, in general, theories have a sample of successes that is large enough. We could think of the confirmation of a new theory as occurring in three stages (this is, of course, schematic):

1. Several candidate theories are proposed in order to account for some phenomena. All are good explanations.
2. A selection occurs on the basis of experiments designed to test the candidates (a "fierce competition", which can involve novel predictions), and only the theory that passes all the tests is retained (this is part of van Fraassen's explanation).
3. The winning theory has a good sample of models that are empirically successful (even more so if it was designed to account for the empirical successes of its predecessors), and it continues to be successfully extended to new domains.

In stage two, we already have a sample of successful models from which we can infer, by induction, that other models of the same theory are also successful. The more numerous the successes during stage three, the more future successes are to be expected. So, we can justify empirical adequacy by induction.

Contrarily to the case of simple induction, our samples are not situations, but models, and the property that is predicated is model adequacy (empirical success for all contexts). These models are representative of their population because they respect the meta-norms of representation associated with the theory, for example in physics, its fundamental laws.⁹ The conclusion of the induction is that all the models of the theory are likely to be adequate, which is to say, per our definition from Sect. 4.3.3, that the theory itself is likely to be adequate. We can reinforce this conclusion by recalling that there is a certain flexibility in model construction. A theory, understood as a family of models, can evolve to accommodate new phenomena, which takes the form of an integration of new models respecting the laws of the theory and incorporating domain-specific postulates. This makes the successful extension of the theory even more likely, because it only has to have one adequate model for new situations (assuming that the uniformity of the theory can be preserved, that is, avoiding ad-hoc postulates: see Sect. 4.3.3). The conclusion of the inference is not that all models respecting the laws of the theory and constraints of coherence are adequate, but that they are more likely to be adequate, so that in general, it will be possible to account for new phenomena by adapting the theory.

This kind of induction also justifies the fact that theoretical models will continue to be successful at higher levels of precision. The problem of underdetermination that affected us previously, due to the fact that many curves can fit the same data points, is resolved, because theoretical selection eliminates all the competitors of the model. Other curves are not as well justified, because they are not part of an overarching theory that is successful in various contexts. For this reason, it is rational to expect that the model will continue to be empirically adequate with higher levels of precision, and only a whole new theory can cast doubt on this inference.

⁹I say a bit more about representativeness in Sect. 6.4.3.

Concerning the need, mentioned in the previous section, of accounting for the success of conjunctions of successful theories, I am confident that a similar rationale can be given (although an analysis of concrete examples would be advisable, instead of the abstract framing usually found in the literature). Note that a simple conjunction of successful theories is not necessarily always successful. As van Fraassen (1980, ch. 4.3) notes, combining theories generally requires correcting them. These corrections require empirical verification before being accepted. However, given the remarks previously made on the fact that unification can be empirically confirmed, it should not come as a surprise that a theory that unifies two successful theories turns out to be successful as well. The unifying theory already accounts for the range of phenomena of the theories that it unifies, and by induction, we can expect it to be successful in new domains. This rationale could also justify the success of the combination of simple models into complex ones.

We have successfully extended van Fraassen's explanation: the "fierce competition" to which theories are submitted entails that the theories we retain are successful for a range of phenomena, and by induction on these first successes, we can infer that it is likely that these theories will be able to make new predictions and to be successfully extended to unprecedented levels of precision.

However, I should say more about induction on models, because this is an induction of a peculiar kind.

6.4 Objections to the Inductivist Response

In this section, I wish to examine some challenges to the proposal of the previous section. They have to do with the fact that the induction involved is not an induction on worldly entities, but an induction on representational entities: theoretical models. This is a second-order induction, so to speak, or what is often called a *meta-induction*.

6.4.1 Sample, Population and Modalities

A first issue with regard to induction on models concerns the right delimitation of the class of entities on which the induction operates. Do they constitute a legitimate population?

A first worry is that there can be latency in the identification of models. A model can be more or less abstract: some dynamical parameter can be fixed, and we get a new model. I do not think that this is a problem for the kind of induction proposed here. One could assume, as Giere (1999, ch. 6) does, that there is a privileged level of abstraction in representation, corresponding to the models that are sufficiently abstract to be useful in various contexts, and sufficiently concrete for the various potential targets of the model to share a high degree of similarity. This could be

associated with the coherence of norms of relevance: the privileged level would be the highest level at which norms of relevance for model application can be made coherent. Alternatively, one could consider that the induction on models operates at various levels of abstraction in a hierarchical way, so as to provide a finer picture of what is or is not justified in a theory.

There is a more serious difficulty, however. Following the account of scientific representation proposed in Chap. 3, general models are abstract entities constituted of a structure and norms of application (Sect. 3.3.3). The properties that are predicated of them in our induction concern their relation to worldly phenomena: their conditions of accuracy are satisfied by the situations to which these models are applicable. These models seem to constitute a legitimate population for inductive reasoning because they are all models of the theory, which means that they all respect the same theoretical laws and principles (Sect. 3.4.1). This is the sense in which these models can be considered representative of their population. However, theoretical models also incorporate their own norms of application, associated with factual knowledge and domain-specific postulates, which determine the set of situations that they are apt to represent. For example, it is part of our factual knowledge that a model of the solar system should incorporate eight planets.

The problem is the following. Given a model, we cannot necessarily know *a priori* which situations it could represent, or if it can represent any actual situation at all. However, it does not seem to make much sense to infer, by induction, that all models that respect the laws of the theory are adequate if some of these models are not apt to represent anything in the universe. The predicate “either being adequate or not representing anything” is not clearly projectible. So, we should restrict our population to the models that do have targets of representation. However, then, our inductive reasoning is problematic, because we are not in a position to specify which models are part of the population by means of well-defined criteria. A model can only be known to be part of the population in retrospect, once it has been constructed and tested. In this context, it is not certain that we can account for novel predictions: assuming, by induction, that the model used for novel predictions had to be empirically adequate before it was actually applied requires assuming that the model was already apt to represent something in the world, which was not certain.

A possibility could be to perform a simple, first-order induction on situations instead of a meta-induction on models: all situations of the world are such that the models of the theory that can represent them are accurate. There would be no problem if general conditions of accuracy for the theory were given once and for all at the theory level. This idea could be based on Craig (1956)’s approach in the context of a syntactic conception of theories (Sect. 2.2): Craig showed that it is always possible to construct a sentence that expresses all and only the empirical consequences of an axiomatic theory. Perhaps every situation of a very generic type, corresponding to being in the domain of application of the theory, satisfies the Craig sentence of the theory, and this could be justified by a simple induction on situations of this type. If this were practicable, the empirical adequacy of theories would be justified in the same way as that of models, as if theories were merely very abstract models. However, invoking Craig’s theorem means moving away from

a model-based approach towards theories, assuming a strict distinction between theoretical and observational vocabulary, and this is idealistic, precisely because the conditions of application of the theory to new phenomena are not necessarily known in advance. If accepting the latter point, the scope of situations on which we can project empirical adequacy would be the ones for which conditions of application of the theory have already been identified, but then, again, the induction could be too limited to account for novel predictions. So, an induction on situations instead of models is not a viable solution.

Let us go back to our initial meta-induction on models. The fact that the relevant population is ill-defined, because we do not know which models apply or not in the world, is a serious problem for a non-modal version of empiricism. However, a modal version of empiricism does have the resources to address it. Take the model of the solar system with Neptune. Prior to the observation of Neptune, scientists did not know whether the model was apt to represent the solar system, so it was not clear whether the model should be part of the population on which the induction on models is performed. However, a modal empiricist can say that this model applies to some possible situations, even if it does not apply to the solar system, so it should be part of the population of models on which the induction is performed. The conclusion of the induction is that the model would be accurate in these possible situations, and as it happens, the solar system is one of them.

Scientists sometimes perform experiments in order to test intriguing consequences of the theory. An example of this is the implementation of an EPR-type experiment by Aspect et al. (1982). Before that project was formed, it could have been unclear to a non-modal empiricist whether or not the adequacy of the corresponding model should be inferred, because perhaps the model did not represent anything in the world. For the modal empiricist, it is clear that it can be inferred, in so far as there are alternative ways actual situations are that can be represented by the model of an EPR-type experiment (or perhaps sub-parts of alternative situations: the ones where the experiment is performed by scientists, see Sect. 4.2.5). There is no need to identify worldly situations that an EPR-type model could represent, so long as we are confident that the model does represent a possible situation.

The assumption that these situations are indeed possible is not always reliable. As explained in Sect. 4.2.5, it can often be considered plausible that a model representing a type of situation that was never encountered before does apply to a possible situation, and the more similar this hypothetical situation is to already encountered situations, the more plausible it is. However, this kind of inference is less secure than the induction involved in justifying the adequacy of a single model: it is conditioned on the principled realisability of experimental situations that the model would be apt to represent. Experimentation can be full of surprises. However, we can be confident that a range of models sufficiently similar to the ones that have already been tested would be applicable, and by induction on models, we can infer that these models would make correct predictions.

We can see that the modal empiricist has an important advantage over non-modal empiricists when it comes to addressing the no-miracle argument. Novel

predictions could be considered possibilities prior to their realisations, at least assuming that the corresponding experiment was realisable, and their successes were therefore warranted by induction, which means that they are no miracle. Since novel predictions are often brought about by controlled interventions, this is related to the considerations made in Sect. 4.4.3 with regard to the interventionist component of experimentation.

6.4.2 *The Base-Rate Fallacy and Fallibilism*

Another objection against meta-inductive arguments of the kind proposed in the previous section is provided by Magnus and Callender (2004). Their objection is directed against the no-miracle argument, but it could apply here as well. They argue that inferring theoretical truth from empirical success is a base-rate fallacy. Let us sketch a version of their objection for the case of empirical adequacy. In a Bayesian framework,¹⁰ the probability that a theory is empirically adequate (EA) given that it is successful (S) is:

$$p(EA|S) = \frac{p(S|EA)p(EA)}{p(S)}$$

Let us assume that the probability of success given adequacy $p(S|EA)$ is equal to 1. The remaining term $p(EA)/p(S)$ is the ratio of empirically adequate theories among the theories that would be successful for the phenomena that we have observed so far. The argument is that this ratio is unknown (we cannot really count all possible theories), and that it could be very low, so that we are not justified in believing that our theories are empirically adequate.

The argument has more force against realism, because truth and success are not homogeneous properties. Evaluating the ratio of true theories among successful ones would require mysterious abilities that, arguably, we cannot have if we only access truth by means of empirical success (note that the objection targets the naive version of the no-miracle argument). In contrast, empirical adequacy is just ideal success, so we have a more standard form of induction. We have a sample of models of a theory, and they are all adequate, so the question boils down to whether this sample is representative of its population, all models of the theory, or whether there is a bias in the way it was selected.

The models of our sample are selected on the basis of the types of situations on which scientists choose to experiment. I do not think that scientists can be blamed for only choosing experimental situations where the theory is likely to succeed. However, there could be a bias due to our situation in the universe and our technological capacities. Take for example Newtonian mechanics: we know,

¹⁰See footnote 14 of the previous chapter on the use of this framework.

in light of relativity theory, that it is adequate in flat spacetime (away from very massive bodies). As it happens, we live in flat spacetime regions, and almost all the situations of the universe to which we have easy access are located in flat spacetime. Before relativity theory was proposed, scientists had almost only tested models of Newtonian mechanics whose targets of representation were located in flat spacetime regions (with the exception of Mercury). So, the impressive history of success of Newtonian mechanics prior to relativity theory was biased in favour of a particular class of models (the ones without very massive bodies) that happen to have a much higher success rate than other models, and they were selected for contingent reasons.

Looking at the history of science, we can see a pattern: initially successful theories can usually be extended to new domains that are not too different from the ones where they have proved successful, up to a point where they fail and a new theory is needed. If we were to associate these with meta-inductions, one meta-induction would be optimistic and the other one pessimistic, and interestingly, this corresponds to the two main arguments in the epistemic debate on scientific realism: the no-miracle argument builds on the observation that extensions of theories usually work, at least up to a point, so it is somehow analog to an optimistic meta-induction, and the pessimistic meta-induction, which will be examined in the next chapter, rests on the observation that eventually, such extensions fail and theories get replaced by new ones.

What lesson shall we draw from this? I would say that this plays in favour of a fallibilist approach. An induction on models is not infallible. We have every reason to believe that our best theories are not absolutely empirically adequate. This is obviously the case because there are many types of phenomena for which scientists have not yet constructed any model. They do not know which model of the theory is relevant. But we can be confident that our theories will evolve and adapt to these new domains. However, we also have every reason to believe that there are domains of experience for which no adequate model can be found, because these domains lie beyond the capacities of extension of the theory. A new theory would be needed to account for them. We can already delimit some of these domains for certain theories: we already know that quantum mechanics and the theory of relativity are incompatible in some circumstances that are presently beyond the reach of our technological capacities, and that a theory of quantum gravity would be required to account for the corresponding domains. Nevertheless, it is still rational to assume that our best theories can be extended to a reasonable range of phenomena that are not too remote from the ones for which they have proved successful.

The eventual failure of the extension of theories is much less dramatic for an empiricist than it is for a realist. Discovering that the scope of empirical adequacy for a theory is limited does not amount to rejecting the claim that this theory is empirically adequate within a limited domain. In general, this domain can be better circumscribed in light of its successor. It is thus possible to amend the theory by restricting its relevance function, so as to maintain that the old theory is still empirically adequate, and still useful. Newtonian mechanics, although superseded by more “fundamental” theories, is still used and developed in a large range of

applications. Furthermore, there is a clear notion of progress, since the domain of adequacy of new theories is in general larger than that of their predecessors.

On the other hand, claiming that a theory is true only in a domain is much more problematic, at least assuming a robust notion of truth such as correspondence truth. For example, claiming that there are forces of gravitation far away from very massive bodies, but not near them, seems to contradict the claim that deformations of spacetime explain gravitational phenomena *everywhere*: in light of general relativity, Newtonian forces of gravitation appear to be superfluous in our ontology. The realist could claim that a certain approximate notion of gravitational force that is derivative from space-time deformations corresponds to “emergent” entities that actually exist far away from massive bodies. However, even assuming that this kind of “approximate existence” makes sense, this precise notion is only available retrospectively. So, the implication seems to be that we do not have a precise notion of what our current theoretical terms actually refer to: our current concepts are probably ill-defined or not defined at all, or only empirically (see Nola 1980).

To sum up, the base-rate fallacy objection is valid, because we have no reason to think that our sample is not biased. However, this only means that we should not pretend that we have already reached universal empirical adequacy, and we can still be confident that our theories are still empirically adequate for a rather large domain of experience.

I will return to the problem of theory change in the next chapter, and examine the solutions that the realists have proposed. However, let us first examine a final objection against induction on models.

6.4.3 *Uniformity of Nature and Relativity*

A final challenge for the induction on models can be presented by examining what simple induction must assume. Hume talked about the uniformity of nature. Goodman (1954) talked about the projectibility of predicates. We have seen in Sect. 5.4.1 that in a Bayesian framework, an assumption of uniformity can take the form of prior probabilities that give equal weights to distributions (for projectible predicates) instead of sequences. In sum, induction is warranted assuming that there are regularities in nature for properties that are denoted by a privileged class of predicates. These regularities ground the notion of representativeness that was used in this work.

Now let us consider the induction on models: what could ground the assumption that theoretical models are representative of other models of the same theory? What kind of uniformity, or what kind of projectibility is at stake? This uniformity or projectibility cannot simply concern the observable features of situations that allow scientists to categorise them into types and confirm their predictions, because a theory implies more than simple regularities. It unifies the regularities observed in various types of objects and various configurations. What unites theoretical models is a shared theoretical vocabulary, and a set of laws and principles expressed in this

vocabulary. We could be tempted to say that what an induction on models implies is not simply that there are regularities in nature, but rather that there is some kind of structure of second-order regularities.

Another way of formulating this point is the following. Simple induction is warranted assuming that there is a right distribution for a class of phenomena, and the aim of inductive inference is to find the right one (for example, if a certain proportion of swans are white, we can infer this proportion by induction). Similarly, an induction on models is warranted assuming that there is a right *theory* capable of unifying a variety of observable phenomena, and the aim of inductive inference is to find the right one.

A theory classifies situations by means of its vocabulary, and constrains the possible states for these situations. Arguably, the way the theory categorises phenomena by means of theoretical posits, for example, the way elements are categorised in chemistry by their atomic number, must be projectible for an induction on models to be warranted. Just as we have to assume that white is a more relevant property than blite, we have to assume that atomic number is a more relevant property than others: elements to which we associate the same atomic number will display similar observable regularities, or modal characteristics for a modal empiricist. Not only that, but the connection between atomic number and other properties will be similar for other atomic numbers. The regularities associated with various atomic numbers are connected in a unified way. A simple induction can confirm a simple regularity, and an induction on models can confirm that the connections between them are correctly accounted for by the theory.

At this point, it would seem that modal empiricism is just a version of standard scientific realism, because the projectibility of the theoretical predicates by which a theory classifies phenomena of interest would imply that these predicates “carve nature at its joints”, which is to say that these theoretical predicates refer to natural properties. Since the theory is empirically adequate in a modal sense, one could say that it attributes shared modal characteristics to the “natural kinds” classified by the theory, so we have something that superficially looks like a dispositional essentialism of the kind proposed by Bird (2007), for instance (this argument is related to Boyd (1985)’s contention that projectibility is theory-laden). According to dispositional essentialism, our theories describe the dispositions associated with natural kinds, which are modal characteristics (a “causal profile”) necessarily possessed by exemplars of these natural kinds.

However, crucial differences remain between modal empiricism and scientific realism, even assuming that all the aspects mentioned above are required for the induction on models to be warranted.

First, I assume that the projectibility of theoretical predicates that are directly connected, in experimentation, to the phenomena (which must be established *before* induction) is itself justified by an induction on experience rather than by an inference to the best explanation. I assume that it is the role of norms of experimentation (which are *not* part of a theoretical models, see Sect. 3.4.2) to ensure this projectibility. To the extent that there would be an interplay between a new theory and its operationalisation for new theoretical terms that are not

directly connected to experience, an induction on models would confirm both the projectibility of these new theoretical terms and the adequacy of the theory. This is not a problem in so far as the new theory and its operationalisation rest on past theories, theoretical terms and experimental practices that have already proved adequate or projectible. If there is circularity, it is never complete. Giving a full account of this process is beyond the scope of this work, but one can have a look at Chang (2004)'s treatment of the establishment of temperature scales as an example of this process.

The fact that the theoretical terms of successful theories are projectible ultimately rests on the stability and coherence of their various operationalisations, or that of the terms to which they are related by theoretical principles, and on the prior projectibility of the predicates of past theories and natural language. If this was enough to claim that theoretical terms “carve nature at its joints”, or refer to natural properties, then at least we would have a justification of realism that does not imply problematic inferences to the best explanation, and that would be an improvement. However, there is no reason to infer as much. As mentioned earlier for Newtonian forces of gravitation, the validity of a concept could be restricted to a range of situations. Furthermore, even if we can extend this range of situations by exploring more contexts, the content of a concept ultimately bears on the way it constrains manipulations and expectations concerning our observations in various contexts, so theoretical concepts remain relative to our technical and cognitive abilities and to our epistemic position. After all, many terms of natural language, such as colour terms, are projectible, but they do not correspond to objective categories of nature that would be independent of our cognitive abilities. Finally, the fact that theoretical terms are interpreted in terms of possible *manipulations*, and not only possible observations, reinforces the idea that they are not merely descriptive.

The same arguments apply to the assumption that our theories correctly unify various regularities. This is justified by induction, not by inference to the best explanation, and there is no reason to assume that this unification gives us the laws of nature, however these are interpreted. The same reasons can be invoked: the validity of the structure of modal regularities given by the theory is presumably restricted to a limited domain, it is expressed in terms of relations between theoretical terms, which are ultimately relative to our epistemic position, and this structure does not appear to be merely descriptive, because it is interpreted in terms of possible manipulations as well as observations. Recall that the notion of necessity associated with possible situations is relative (see Sect. 5.2.2). So, there is no reason to assume that the structure of theories correspond to mind-independent laws of nature or to dispositional essences associated with an absolute notion of necessity.¹¹

¹¹These remarks are particularly relevant for addressing the case of structural realism, which is the position according to which our scientific theories correctly describe the structure of the world. Arguments regarding domain limitation will be developed in the next chapter, so as to explain why modal empiricism is not a version of structural realism. I will return to the idea that theories are not merely descriptive in Chap. 8.

In sum, an induction on models is apt to account for the successful extension of theories to new domains of experience without having to assume that our theories, when interpreted realistically, are true. This should be clear, given that nothing in the structure and norms of application of theoretical models implies that one realist interpretation or the other (in the sense given in the introduction of this chapter) must be true. This conclusion is reinforced by the fact that the space of possible theories is much larger, and much less graspable, than the space of possible distributions for a simple property, so that an induction on models is much less secure than a simple induction. This means that many theories could account for the phenomena observed so far without being indefinitely extensible to new domains, which supports the fallibilist approach mentioned earlier.

The main difference between induction and inference to the best explanation is that induction does not allow us to reach a new ontological level that would lie beyond the level on which the induction is performed: the inference bears on the same kind of entities and properties as the ones that are present in our initial sample. According to modal empiricism, theoretical terms and laws are nothing but elaborate systematisations of possible observations and manipulations, and empirical adequacy is nothing but a norm of ideal success for these systematisations. For this reason, even assuming that an induction on models, which is required to justify empirical adequacy, presupposes a stronger form of uniformity of nature than simple induction does, a uniformity that allows such systematisations to operate successfully, it is still insufficient to reach realist conclusions.

6.5 How Far Are We From Realism?

We have seen that the main line of defence for realist positions consists in assuming that the non-empirical criteria that make theories good explanations of the phenomena they predict are indicators of truth. This assumption is itself justified by an inference to the best explanation: this is what explains the impressive empirical success of science, and in particular, the successful extension of theories to new domains of experience. However, there are many reasons to doubt that non-empirical criteria are truth-conducive. They are more likely to be pragmatic or strategic, or desirable for their own sake if the aim of science is to provide unified theoretical frameworks that are empirically adequate.

We have seen that justifying empirical adequacy, including the capacity of theories to make novel predictions, does not require assuming that theories are true, so long as we accept the validity of a particular form of induction: an induction on models. Only the theories that are initially more successful than their competitors are retained by scientists, and since these theories already have a record of successes for a variety of their models, one can expect, by induction, that their other models will be successful as well.

This kind of induction is peculiar. If simple induction presupposes the uniformity of nature and the projectibility of observational predicates, this one presupposes a unified structure, and the projectibility of theoretical predicates. It could seem, then, that modal empiricism is not very different from standard realism. However, there is a crucial difference, which lies in two related assumptions: first, that the projectibility and success of a theory is not unrestricted, but limited to a domain (which, as we have seen, is problematic for a realist), and secondly, that the content of our theories is not purely descriptive, but relative to a perspective on reality associated with our position as active epistemic agents.

The fact that theoretical terms are projectible and that theories correctly describe the structure of modal regularities between possible manipulations and observations, as categorised by these theoretical terms, all this *within a domain*, is enough to save most of the intuitions that motivate realism. It gives us an ersatz version of realism, so to speak. Modal empiricism has all the surface features of standard realism. It makes no real difference with regard to what we could *practically* expect from our scientific theories, at least not until the limits of applicability of these theories is reached. However, it is not a realist position in the standard sense of the term, which would include its semantic component, because no metaphysical conclusions are drawn from these surface features.

In the next chapter, I will give more details concerning these differences, by comparing modal empiricism to the closest version of realism: structural realism. I will argue that these differences play a crucial role in addressing the problem of theory change that affects scientific realism.

References

- Acuña, P. & Dieks, D. (2014). Another look at empirical equivalence and underdetermination of theory choice. *European Journal for Philosophy of Science*, 4(2), 153–180. <https://doi.org/10.1007/s13194-013-0080-3>.
- Aspect, A., Dalibard, J., & Roger, G. (1982). Experimental test of Bell's inequalities using time-varying analyzers. *Physical Review Letters*, 49(25), 1804–1807. <https://doi.org/10.1103/PhysRevLett.49.1804>.
- Belot, G. (2012). Quantum states for primitive ontologists. *European Journal for Philosophy of Science*, 2(1), 67–83.
- Bird, A. (2007). *Nature's metaphysics: Laws and properties*. Oxford: Oxford University Press.
- Bokulich, A. (2016). Fiction as a vehicle for truth: Moving beyond the ontic conception. *The Monist*, 99(3), 260–279. <https://doi.org/10.1093/monist/onw004>.
- Boyd, R. (1980). Scientific realism and naturalistic epistemology. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, 1980*, 613–662.
- Boyd, R. N. (1985). The Logician's dilemma: Deductive logic, inductive inference and logical empiricism. *Erkenntnis*, 22(1-3), 197–252. <https://doi.org/10.1007/BF00269968>.
- Cartwright, N. (1983). *How the laws of physics lie* (vol 34). Oxford: Oxford University Press.
- Chakravartty, A. (2008). What you don't know can't hurt you: Realism and the unconceived. *Philosophical Studies*, 137(1), 149–158. <https://doi.org/10.1007/s11098-007-9173-1>.
- Chang, H. (2004). *Inventing temperature: Measurement and scientific progress*. Oxford: Oup USA.

- Craig, W. (1956). Replacement of auxiliary expressions. *The Philosophical Review*, 65(1), 38. <https://doi.org/10.2307/2182187>.
- Devitt, M. (2011). Are unconceived alternatives a problem for scientific realism? *Journal for General Philosophy of Science/Zeitschrift für Allgemeine Wissenschaftstheorie*, 42(2), 285–293.
- Earman, J. (1993). Underdetermination, realism, and reason. *Midwest Studies in Philosophy*, 18(1), 19–38.
- Forster, M., & Sober, E. (1994). How to tell when simpler, more unified, or less *Ad Hoc* theories will provide more accurate predictions. *The British Journal for the Philosophy of Science*, 45(1), 1–35. <https://doi.org/10.1093/bjps/45.1.1>.
- French, S. (2014). *The structure of the world: Metaphysics and representation*. Oxford, United Kingdom: Oxford University Press.
- Ghins, M. (2017). Defending scientific realism without relying on inference to the best explanation. *Axiomathes*, 27(6), 635–651.
- Giere, R. (1999). Science without laws. *Science and its conceptual foundations*. Chicago: University of Chicago Press.
- Godfrey-Smith, P. (2008). Recurrent transient underdetermination and the glass half full. *Philosophical Studies*, 137(1), 141–148. <https://doi.org/10.1007/s11098-007-9172-2>.
- Goodman, N. (1954). *Fact, fiction & forecast* (vol. 25). London: University of London.
- Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science* (vol. 22). Cambridge: Cambridge University Press.
- Hitchcock, C., & Sober, E. (2004). Prediction versus accommodation and the risk of overfitting. *The British Journal for the Philosophy of Science*, 55(1), 1–34. <https://doi.org/10.1093/bjps/55.1.1>.
- Kukla, A. (1996). Does every theory have empirically equivalent rivals? *Erkenntnis*, 44(2), 137–166. <https://doi.org/10.1007/BF00166499>.
- Laudan, L., & Leplin, J. (1991). Empirical equivalence and underdetermination. *Journal of Philosophy*, 88(9), 449–472.
- Leplin, J., & Laudan, L. (1993). Determination underdetermined: Reply to kukla. *Analysis*, 53(1), 8–16. <https://doi.org/10.1093/analys/53.1.8>.
- Levin, M. (1984). What kind of explanation is truth. In J. Leplin (Ed.) *Scientific realism* (pp. 124–139). Berkeley: University of California.
- Lipton, P. (2004). Inference to the best explanation (2nd edn.). *International library of philosophy*. London; New York: Routledge/Taylor and Francis Group.
- Lyons, T. (2002). Explaining the success of a scientific theory. *Philosophy of Science*, 70(5), 891–901.
- Magnus, P., & Callender, C. (2004). Realist ennui and the base rate fallacy. *Philosophy of Science*, 71, 320–338.
- Maxwell, G. (1970). Theories, perception and structural realism. In R. Colodny (Ed.) *The nature and function of scientific theories* (pp. 3–34). Pittsburgh: University of Pittsburgh Press.
- McAllister, J. W. (1989). Truth and beauty in scientific reason. *Synthese*, 78(1), 25–51. <https://doi.org/10.1007/BF00869680>
- McMullin, E. (1984). A case for scientific realism. In J. Leplin (Ed.) *Scientific realism* (pp. 8–40). Berkeley: University of California.
- Myrvold, W. C. (2003). A Bayesian account of the virtue of unification. *Philosophy of Science*, 70(2), 399–423. <https://doi.org/10.1086/375475>.
- Myrvold, W. C. (2016). On the evidential import of unification. *Philosophy of Science*, 84, 92–114.
- Nola, R. (1980). Fixing the reference of theoretical terms. *Philosophy of Science*, 47(4), 505–531.
- Norton, J. (2008). Must evidence underdetermine theory? In M. Carrier, D. Howard, J. Kourany (Eds.) *The challenge of the social and the pressure of practice: Science and values revisited* (vol. 14, pp. 17–44). Pittsburgh: University of Pittsburgh Press.
- Nyrup, R. (2015). How explanatory reasoning justifies pursuit: A peircean view of IBE. *Philosophy of Science*, 82(5), 749–760.
- Okasha, S. (1997). Laudan and Leplin on empirical equivalence. *The British Journal for the Philosophy of Science*, 48(2), 251–256. <https://doi.org/10.1093/bjps/48.2.251>.

- Peirce, C. S. (1931). *Collected papers of Charles Sanders Peirce*. Cambridge: Harvard University Press.
- Psillos, S. (1999). Scientific realism: how science tracks truth. *Philosophical issues in science*. London; New York: Routledge.
- Putnam, H. (1975). The meaning of 'Meaning'. *Minnesota Studies in the Philosophy of Science*, 7, 131–193.
- Salmon, W. (1990). Rationality and objectivity in science or Tom Kuhn meets Tom Bayes. In C. W. Savage (Ed.) *Scientific theories* (pp. 14–175). Minnesota: University of Minnesota Press.
- Smart, J. (1963). *Philosophy and scientific realism*. London: Humanities Press.
- Stanford, K. (2017). Underdetermination of scientific theory. In E. Zalta (Ed.) *The Stanford encyclopedia of philosophy*, winter 2017 edn. Stanford: Metaphysics Research Lab, Stanford University. <http://plato.stanford.edu/archives/win2009/entries/scientific-underdetermination/>.
- Stanford, P. K. (2010). *Exceeding our grasp: Science, history, and the problem of unconceived alternatives*. New York; Oxford: Oxford University Press.
- van Dyck, M. (2007). Constructive empiricism and the argument from underdetermination. In B. Monton (Ed.) *Images of empiricism* (pp. 11–31). Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199218844.003.0002>.
- van Fraassen, B. (1980). *The scientific image*. Oxford: Oxford University Press.
- van Fraassen, B. (1989). *Laws and symmetry* (vol. 102). Oxford: Oxford University Press.
- Zahar, E. (1973). Why did Einstein's programme supersede Lorentz's? (II). *The British Journal for the Philosophy of Science*, 24(3), 223–262. <https://doi.org/10.1093/bjps/24.3.223>.

Chapter 7

Theory Change



Abstract The induction on models presented in the previous chapter brings modal empiricism close to structural realism in some respects. In this chapter, I examine the differences between the two. I review the main motivation for structural realism, which is to account for a continuity in theory change, and the main objection against structural realism: Newman's objection. Modalities are invoked by structural realists as a solution to the problem. I show that they are confronted with a dilemma: assuming too strong a notion of modality, the position cannot account for continuity in theory change, and assuming a notion that is weak enough, the position cannot be qualified as realist. This reveals the two main differences between modal empiricism and structural realism: first, the fact that empirical success is not explained by appealing to real entities, and second, the fact that modalities are relative to our epistemic position. The conclusion is that modal empiricism is not structural realism, and that it is better equipped to confront the problem of theory change.

7.1 The Pessimistic Meta-Induction

I have explained in the previous chapter how modal empiricism can account for the empirical success of scientific theories. Scientific realism resorts to theoretical truth to explain this success. I have already presented various arguments against this idea, but I now wish to turn my attention to what is often considered the most serious one: the so-called pessimistic meta-induction (Laudan 1981). In this chapter, I will examine how modal empiricism fares with respect to this argument in comparison with various strands of realism, and in particular, in comparison with a position that was specifically designed to respond to this argument: structural realism.

The pessimistic meta-induction can be seen as a direct refutation of the no-miracle argument for scientific realism presented in the previous chapter. Remember that according to the no-miracle argument, scientific realism is the only philosophy that does not make the empirical success of science a miracle. Theories work because they are true or approximately true. Theoretical terms refer to natural properties, and theoretical structures correctly describe the relations between these

properties. However, there seem to be many counterexamples to this idea: many theories of the past were very successful, including for making novel predictions, and yet they were eventually abandoned and replaced by better ones. This would not be a problem if the new theories were mere extensions of the past ones, but this is not the case: there is no ontological continuity between them. Scientists once postulated that a caloric fluid was responsible for heat phenomena, that phlogiston was responsible for combustion phenomena, and that light was constituted of waves propagating in the ether, but today, scientists do not consider that “caloric”, “phlogiston” and “ether” refer to anything in the world. These counterexamples are enough to undermine the no-miracle argument, but we can even go further: by induction on past theories, we have good reasons to think that our current theories will eventually be abandoned and replaced by better ones, and that their theoretical terms do not refer to anything in the world. This idea is related to Stanford (2010)’s argument that current theories are probably underdetermined, because unconceived alternatives could account for the same phenomena (Sect. 6.2.1).

The standard realist responses to the pessimistic meta-induction consist in limiting realist claims. For example, theoretical truth can be restricted to some theories only (“mature” theories, the ones that allowed for novel, unexpected predictions), and to parts of these theories (the components that were essentially involved in novel predictions) (Psillos 1999). Stanford (2010) argues that this move is only available in retrospect. Lyons (2017) also argues against this strategy by giving numerous examples of successes on the basis of false posits in the history of science. In response, one could understand realism as a “stance” towards science (in van Fraassen 2002’s sense) that should be applied on a case-by-case basis, rather than as a general theory about science (Saatsi 2017). This would mean adopting a piecemeal approach instead of attempting to provide general recipes for extracting realist commitments from scientific theories, and this would be a better strategy for dealing with counterexamples. It is hard to argue against stances, so I will not discuss this proposal here, but it should be clear that the stance adopted in this book is different.

More radical departures from standard realism that still purport to give us “general recipes” for extracting realist commitments consist in restricting truth to some aspects of theories only. This is the case for entity realism, according to which we can be realist about postulated entities in so far as we can causally interact with them (Cartwright 1983; Hacking 1983). Entity realists focus on local causal relations rather than on laws of nature, thus rejecting the idea that the structure of theories corresponds to the nomological structure of reality. Furthermore, they do not necessarily resort to an inference to the best explanation to justify their position, but rather to causal reasoning. The fact that experimental techniques generally survive theory change supports the idea that the pessimistic meta-induction does not affect this kind of realism. Note, however, that Herschel thought he had isolated caloric with a prism (today we would say that it was infrared radiations) (Chang and Leonelli 2005), so one could wonder if causal interactions are enough for being a realist.

One criticism of entity realism (which is a bit of a caricature) is that we cannot be committed to entities without entertaining beliefs about them, so that believing in the existence of theoretical entities without being a realist about theories would be incoherent. More subtly, it can be argued, on the basis of historical cases, that the stabilisation of background theories generally participates in the consolidation of our beliefs that real entities are being manipulated in experiments, and sometimes (for example in astronomy), manipulations are not even required (Morrison 1990). This might undermine entity realism as a faithful account of scientists' inferences, but this does not make the position incoherent. Entity realists do not deny that theories serve as reliable guides for local causal interactions with theoretical entities, and this does not require being a realist about the full theories. Cartwright is realist about phenomenological laws, for instance, but not about fundamental laws (see Clarke 2001).

I am sympathetic to the arguments of entity realists, but I do not think that they warrant a realist semantics, according to which theoretical terms directly refer to real entities. Scientific models have a modal structure which relates possible observations and manipulations, and this might be enough to interpret them in terms of a causal structure (for example, assuming a counterfactual theory of causation, and assuming that models are empirically adequate in the modal sense). This causal structure is vindicated by our successful interactions with concrete targets. It might also be convenient to put names on the nodes of this causal structure. However, I would not draw too many ontological conclusions from this. I suspect that the stability of our causal interactions with reality is in general relative to background conditions and to intentional attitudes, so that the entities that we name do not necessarily exist in an absolute sense, that is, independently of our conceptualisation, activities and contingent position in the world.

Another position that restricts realist claims to one aspect of theories is structural realism. It takes the opposite view to entity realism: postulated entities do not exist (or we should be agnostic about their existence), but the structures of theories correctly describe the nomological structure of reality. This structure is said to be modal, causal, or simply extensional depending on the version.

The term "structural realism" was coined by Maxwell (1970) in reference to a position first proposed by Russell (1927). Drawing inspiration from a similar position proposed by Poincaré (1902), Worrall (1989) has re-introduced structural realism in the contemporary debate in order to address the pessimistic meta-induction specifically. Two versions of this position are often distinguished: an epistemic version and an ontic version. According to the epistemic version, we can know the *structure* of reality, but not its *nature*. So, we should remain agnostic about the nature of the world beyond its structure: we should not interpret theoretical terms. We should not assume, for example, that they refer to natural properties. According to the ontic version (Ladyman and Ross 2007; French 2014), the nature of the world *is* structural: structure is "all there is".

According to its proponents, structural realism can respond to the pessimistic meta-induction because there is a continuity of structure between successive theories: the structure of abandoned theories is generally retained in theory change.

For example, although there is no caloric fluid, the caloric theory correctly described the relations involved in heat phenomena. Therefore, we can claim that past theories, although they were eventually abandoned, are still “structurally true”, or approximately so, and we can be confident that our current theories are also true in this sense. Or so the argument goes. Structural realism also does justice to the no-miracle argument because a correspondence of structure between our theories and the world is enough to explain their empirical success. Thus we can have, to borrow Worrall’s words, the “best of both worlds” between realism and empiricism.

Ontic versions of structural realism are also motivated by their ability to address issues in the metaphysics of contemporary physics, in particular, questions related to the identity of particles in quantum mechanics. The probabilities derived from quantum-mechanical models are more naturally attributed to physical configurations, independently of the objects that realise these configurations, than to identifiable particles (a permutation of two particles of the same kind is not “counted” as a different configuration for probability attributions). This would make an ontology of particles superfluous, hence, we should adopt a structuralist ontology.

Modal empiricism is not very far from structural realism in some respects. Note that the positions of Poincaré and Russell from which structural realism originates followed a distinct argumentative strategy from the one that contemporary philosophers adopt. It did not consist in weakening scientific realism in order to respond to anti-realist arguments (although Poincaré did provide an argument to this effect), but rather in starting from an empiricist position, and answering the question: what can we know about reality? According to Poincaré, only structural aspects can be communicated, so only structural aspects are susceptible of constituting public knowledge.¹ According to Russell, the relations between the basic units of perception, or sense-data, reflect the relations between unknowable worldly entities, so we can know the form of these real relations, but nothing more. These approaches correspond to what Psillos (2001) calls an upward, as opposed to a downward, path towards structural realism.

In one sense, the present work is the proposal of an upward path towards modal empiricism, which makes the position close to structural realism. Starting from contextual representational activities, the question that I wish to answer is: what can we know, or what can we rationally accept? And the answer is: we can accept that our best scientific theories are modally adequate, in the sense defined in Chap. 4. This upward path was completed in Chaps. 5 and 6, with the suggestion that an induction on situations, contexts and models can justify this notion of empirical adequacy. However, I refrain from referring to the resulting position as a version of structural realism. One of the reasons for this is that there is an influential objection against structural realism called Newman’s objection, which affects all its versions, and I believe that this objection renders structural realism untenable. According to it, structural realism would either be trivial or collapse into empiricism: it is merely committed to relations between our observations. Interestingly enough, it has been

¹Schlick (1932) entertained similar views.

claimed that thinking of these relations as modal would constitute a way out of this conundrum. However, I believe that this move does not work.

In this chapter (which is based on Ruyant 2019), I wish to show precisely how modal empiricism differs from structural realism. I will start from an examination of Newman's objection and the responses that have been proposed, and this will soon lead us back to the pessimistic meta-induction and to the problem of theory change. I will argue that it is impossible to escape Newman's objection without falling prey to the pessimistic meta-induction. Talking about modal relations does not help: either these relations are not "real", or they do not survive theory change. This depends on the interpretation of modalities, and this is exactly where the difference between modal empiricism and structural realism lies. It entails that structural realism is unable to respond to the pessimistic meta-induction, contrarily to what its defenders claim.

7.2 Structural Realism and Newman's Objection

Structural realism purports to be a position of compromise, capable of responding to the pessimistic meta-induction while doing justice to the no-miracle argument. Objections against the position can be broadly classified into two types: the objections to the effect that it cannot answer the pessimistic meta-induction any more than standard realism does, either because it is indistinct from standard realism, or because there is no structural continuity in theory change, and the objections to the effect that it cannot do justice to the no-miracle argument any more than empiricism does. Let us examine these objections.

7.2.1 *The Objections Against Structural Realism*

Regarding the first type of objection, Psillos (1995) claims that the nature versus structure dichotomy employed by structural realists is unclear. Psillos is a realist, and he believes that standard realism has the resources to answer the pessimistic meta-induction, but that structural realism is not a consistent alternative since there is a continuum between nature and structure rather than a strict dichotomy. On the one hand, the structure of a theory informs us about the nature of the entities it posits. For example, how Newtonian mass relates to force and acceleration informs us about the nature of mass. On the other hand, structure must be interpreted for a theory to make any prediction, so structure alone cannot explain empirical success. In a similar vein, Papineau (1996, p. 12) claims that a restriction of realism to structure is no restriction at all. In response, Votsis (2004) notes that structural realism does not say that structure is not interpreted at all, but that it is only interpreted empirically. Frigg and Votsis (2011) propose to explicate the distinction between nature and structure in terms of intension and extension. However, this proposal does not work

for the modal versions of structural realism that will be examined later, because they are not extensional. Alternatively, the emphasis can be put on the distinction between structuralist and object-based ontologies, instead of the distinction between nature and structure (Ladyman and Ross 2007, p. 156–157). This approach fits well with versions of structural realism that are motivated by contemporary physics.

In any case, as noted by Newman (2005), the structural realists must explicate what exactly structural continuity amounts to, beyond the intuitive idea that equations are somehow transposed from one theory to the other. Without any such account, they cannot claim to respond to the pessimistic meta-induction. Some authors argue that there is no real continuity of structure between successive theories, apart from empirical structures. Psillos (1999, ch. 7) claims that not all structures are retained in theory change. Stanford (2003) takes the example of Galton's law in biology, where the structural continuity is only empirical with Mendel's theory, but no relations between posited entities (here the genetic characteristics of ancestors of an organism) are retained. Although they do not take it to undermine scientific realism, Saatsi and Vickers (2011) give the example of Kirchhoff's theory of diffraction, which was empirically successful despite being inconsistent and positing a wrong structure, in light of Maxwell's theory. Redhead (2001) takes the example of the commutativity of observables in classical and quantum mechanics to argue that structural transformations are too important in theory change to really talk about structural continuity (However, see Thébault 2016). Bueno (2008) provides other examples of structural loss in theory change. Pashby (2012) uses the example of the discovery of the positron to show how shifts in the metaphysical commitments of a theory may also be displayed in terms of changes in theoretical structure. This supports Psillos's contention that the nature–structure dichotomy is not so clear.

A common strategy for the structural realist facing this type of difficulty is to concentrate on well-confirmed empirical relations between phenomena that are very unlikely to disappear in theory change. Structural realists can refer to the principle of correspondence (Post 1971, p. 228) for their purpose. According to this principle, “any acceptable new theory *L* should account for the success of its predecessor *S* by ‘degenerating’ into that theory under those conditions under which *S* has been well confirmed by tests”. However, then, the ultimate benchmark for structural continuity between theories is their predictive success, and it is not so clear that anything is retained in theory change beyond empirical structures that are approximate and restricted to a limited domain of experience: an empiricist would readily agree.

This brings us to the second family of objections against structural realism: that it does not say anything beyond what empiricism already says. The most influential objection to that effect is dubbed “Newman's objection”, after Newman (1928)'s argument against Russell's version of structural realism. The argument can be expressed intuitively as follows: if relations are conceived of in purely extensional terms (a relation is defined by the objects it relates), then it suffices that some objects exist for all relations that could be defined between these objects to exist as well, mathematically speaking. So, for any mathematical structure to exist, it is sufficient that there exist enough objects in the world to bear that structure, and

structural realism is a trivial position that is only committed to a cardinal claim (the existence of a certain number of objects in the world). To block this objection, the structuralist must say more, and qualify the kind of relations that she is talking about to differentiate them from purely mathematical relations, but then, the position is not purely structural: the structure must be interpreted.

As already noted, structural realism can accept that the structure be interpreted empirically. If such is the case, this position is only committed to a structure of observations (and a sufficient number of inaccessible objects to bear that structure), and the position is only superficially different from empiricism.

7.2.2 *How to Escape Newman's Objection?*

A useful framework to examine this objection is the Ramsey sentence formalism. It starts from the assumption that the cognitive content of a theory can be expressed as a logical sentence $\prod$ in a vocabulary comprising observational terms O_i and theoretical terms T_i .

$$T = \prod(O_1, O_2, \dots, O_n, T_1, T_2, \dots, T_m)$$

The idea that theoretical terms should not be interpreted can be implemented by replacing theoretical terms by variables t_i over which one quantifies existentially: this is the Ramsey sentence of the theory.

$$(\exists t_1, t_2, \dots, t_m) \prod(O_1, O_2, \dots, O_n, t_1, t_2, \dots, t_m)$$

The difference with the former formulation is that while the T_i , as part of our vocabulary, are interpreted, the t_i are not. According to Maxwell (1971), the commitments of structural realism can thus be expressed by “ramseyfying” theories: only the observational terms are interpreted, but the structure of the theory is retained. In particular, one does not have to assume that the t_i refer to natural properties: they could be multiply realisable in light of a new theory (French 2014, ch. 5.9). This is how structural realism, thus formulated, can answer the pessimistic meta-induction: if these terms do not refer to natural properties, it is not necessarily a problem if they are eventually abandoned. But following Newman's objection, the position cannot do justice to the no-miracle argument, because it does not explain observable regularities: it is a mere synthetic statement of these regularities, plus a cardinal claim.

In this framework, Newman's objection takes the form of a theorem of second-order logic (Demopoulos and Friedman 1985): the Ramsey sentence of a theory is true if and only if all the empirical consequences of the theory (the observational sentences that can be deduced from the theory) are true, and if there exist a certain number of objects.

Some authors, such as Worrall and Zahar (2004), claim that this is sufficient for structural realism. The Ramsey sentence expresses all the empirical consequences of a theory, for all past, present and future phenomena in the universe, in the most compact way possible, and this, in itself, is informative. Note that constructive empiricism is also committed to our theories being empirically adequate for all past, present and future phenomena. Van Fraassen also claims that theories must be informative. If it were not for the Ramsey formalism employed here, a modal empiricist could make a similar claim: theories are conceptual schemes that account for all observable phenomena in a compact (unified) way. So, it is far from clear what “informative” means here, and why this would be enough for realism. It is not obvious that this compact formulation is retained in theory change, in particular if the theoretical terms of past theories no longer appear in new theories. If saying that the formulation is informative is a way to claim that the properties and relations the t_i refer to are “natural” rather than fictitious, then we are back to standard realism, and we have not solved anything.

Being a theorem of second-order logic, Newman’s objection is inescapable if one assumes that a Ramsey sentence is the right way to express the content of a theory to which structural realism is committed. But perhaps this formalism unduly attributes implicit commitments to the structural realist. Among these implicit commitments are (i) the assumption of a dichotomy between observational and theoretical terms, (ii) the assumption that a theory quantifies universally over objects, properties and relations, and (iii) the assumption that the content of a theory is extensional.

Melia and Saatsi (2006) suggest different ways of departing from this formalism to escape Newman’s objection by rejecting one or the other of these assumptions. For example, one can relax assumption (i) by adding mixed predicates to the vocabulary: predicates that apply indifferently to observable or unobservable objects (for example, an atom is smaller than a molecule just as a cat is smaller than a dog). This is the solution envisaged by Cruse (2005).² One can also relax assumption (ii) by restricting the intended domain of the theory instead of assuming universal quantification. This can apply either to objects (for example, the theory applies only to concrete or physical objects, not all objects) or to properties and relations. However, Melia and Saatsi argue convincingly that none of these solutions works. Introducing mixed predicates is too limited a solution. Restricting the domain of objects merely amounts to introducing a new predicate into the theory, and Newman’s objection also applies to the new ramseyfied theory. Finally, restricting the domain of properties and relations, which amounts to introducing a second-order predicate, either brings us back to standard realism if the qualification is too strong, and structural realism cannot respond to the pessimistic meta-induction anymore, or it is not sufficient to block Newman’s objection. For example, invoking qualitative

²It could be associated with transduction, which, according to McGuire (1970), is a mode of inference that played a role in Newton’s atomism. It amounts to extending characteristics of observable objects to hypothetical, unobservable objects.

properties, which would correspond to combinations of natural properties, would lead to triviality, because quite any structure of qualitative properties can be realised.

This leaves us with one option: rejecting assumption (iii) and incorporating non-extensional logic, such as modal logic, to express the fact that theoretical relations express nomological relations rather than mere universal regularities. This would be the sense in which the structure posited by scientific theories is "real". Melia and Saatsi claim (but do not prove) that this would block the theorem supporting Newman's objection. This is part of what motivates many structural realists to assume that the structure of reality that theories describe is a modal structure.

7.2.3 *Transposition to the Semantic View*

This debate is framed in a statement view of theories. One could wonder how it translates to a semantic view, where theories are not construed as statements, but rather as families of models (see Sect. 2.2.2). French and Ladyman (2003) suggested that Newman's objection is an artefact of the statement view. In the context of a semantic view, structural correspondence between theories and the world could be expressed in terms of isomorphism, or partial isomorphism, between the models of the theory and their targets of representation, assuming that the domain of application of various models is specified. This apparently constitutes a more radical departure from the Ramsey sentences formalism than the ones just examined. However, Newman's objection was initially framed independently of the Ramsey sentences formalism. It is quite general (and it is presented as such by Cei and French 2006). Ainsworth (2009) shows that Newman's objection can be transposed to the semantic view. This was later acknowledged by French (2014, p. 126).

This point was already briefly mentioned in Sect. 2.2.3. The general idea is that isomorphism only makes sense if one gives a set of "important" properties and relations that are preserved by isomorphism. In our two-stage account of representation, or in the realist variant sketched in the introduction of Chap. 6, these important properties and relations are the ones that are mapped to properties and relations of the target of representation when the model is interpreted. In order to avoid the problem of theory change, this mapping must not involve the assumption that theoretical terms directly refer to natural properties. This suggests that only observable features of the target are mapped to components of the model: the model is only interpreted empirically. Assuming this, having an isomorphism between the model and the target only means that the model correctly describes the regularities between these observable features, and structural realism is indeed just a version of empiricism. This roughly amounts to opting for our two-stage account of representation instead of its realist variant.

Melia and Saatsi's proposals can easily be transposed to this approach. One could interpret the model by including unobservable entities in the mapping, in so far as the corresponding symbols could also be mapped to observable entities

in other contexts, which is equivalent to introducing a mixed predicate. One could also restrict the domain of quantification by qualifying the “important” properties and relations, which, again, amounts to including them in the mapping. And for the same reasons, this will not do: either too many properties and relations will act as potential truth-makers for the model, which trivialises theoretical truth, or they will not survive theory change, and we cannot answer the pessimistic meta-induction anymore.

An interesting option for the structural realist could consist in focusing on higher-level relations: relations between the models of the theory, instead of the structure of models themselves. These higher-level relations correspond to the laws and principles of the theory, or what I have called meta-norms of representation. This option fits nicely with French (2014)’s approach, and notably, his emphasis on theoretical symmetries. Theoretical symmetries, such as space-time translation symmetries, are transformations of theoretical models that preserve parts of their structures, notably the fact that they still respect the laws of the theory, and these symmetries often serve as a guide for developing new theories. They are sometimes called “meta-laws”. In the standard model of particles in physics, fundamental particle kinds and their interactions are characterised by means of symmetry groups. Ontic versions of structural realism generally claim that theoretical symmetries correspond to the fundamental structure of reality. Since symmetries are characterised in terms of invariance under change, this approach is similar in spirit to the time-honoured association between perspective invariance and objectivity (which one can find in the work of Husserl, for instance).

Could focusing on the structure that relates the models of a theory help escape Newman’s objection? There is not enough space here to analyse this solution in detail, but taking an empiricist stance, I do not see any reason to consider that theoretical symmetries are anything more than syntheses of empirical regularities. There is an ongoing discussion about the empirical status of symmetries (Kosso 2000; Greaves and Wallace 2014). In yet unpublished work (Ruyant 2020), I argue that assuming that scientific models are indexical, and that the aim of theories is not to represent the whole universe, but to successfully apply to bounded situations, all theoretical symmetries have an empirical status, for example, regarding the equivalence of various possible operationalisations on the same object. In this respect, symmetries, contrarily to models, do not seem to have a direct *representational* status, and I think that it is a mistake to reify them. Furthermore, if higher-level relations merely encapsulate what all models of the theory have in common, then the solution offered to Newman’s objection amounts to picking “important” relations (the ones that ensure the empirical stability of various applications). So, we are back to the same problem: the question is to what extent this higher-level structure survives theory change, for example, to what extent the space-time symmetries of Newtonian mechanics can be considered approximations of that of general relativity. The arguments of the next section will be relevant in this respect.

The final solution proposed by Melia and Saatsi is quite similar to adopting modal empiricism. The idea is that a model does not only describe extensional regularities between observable aspects of the target, but also intensional regularities, that is,

regularities between possible observations. In this sense, modal empiricism is very similar to the versions of structural realism that adopt this solution, if not identical. Ladyman (2004) suggests that a version of empiricism committed to modalities would be nothing but structural realism.

However, it remains to be shown that this solution works, and there is an important aspect that remains unspecified, and that plays a crucial role: the interpretation of modalities. Depending on how modalities are interpreted, the structure presented by scientific models could count as real or not, and it could survive theory change or not, and I will argue that it has to be one or the other.

7.3 “Real” Relations and Theory Change

Newman’s objection threatens to make structural realism indistinct from empiricism: the relational structure that theories describe would merely correspond to observational regularities. As we have just seen, the standard response to this problem consists in talking about *modal* relations, as opposed to merely extensional relations. If this solution works, then should modal empiricism not be classified as a version of structural realism? In this section, I will argue that this is not the case.

7.3.1 Newman’s Objection in Modal Logic

In order to make this point, let us see how modal logic affects Newman’s objection when this objection is expressed as a theorem of second-order logic in the Ramsey-sentence formalism. It is well-known that one can mimic modal logic in extensional logic by quantifying over possible worlds. If we follow Melia and Saatsi’s suggestion, we can express the cognitive content of a theory as follows, where the w_i are possible worlds:

$$(\exists t_1, t_2, \dots, t_m)(\forall w_1, w_2, \dots, w_p) \prod (O_1, O_2, \dots, O_n, t_1, t_2, \dots, t_m, \dots, w_1, w_2, \dots, w_p)$$

Perhaps a theory also has existential quantifiers on possible worlds, not just universal quantifiers, to express possibilities. Perhaps quantifiers are embedded into its formulation rather than appear at the beginning. But this does not matter for the sake of the argument. In any case, we can see that this Ramsey sentence is not formally different from the previous one, and there is no reason why the same theorem would not apply. The conclusion of the theorem would be slightly different, though, since possible worlds are interpreted objects that should not be ramseyfied. The conclusion would be the following: the modal Ramsey sentence of a theory is true if and only if all the empirical consequences of the theory are true in all possible worlds, and if there are enough objects in possible worlds to bear the structure of

the theory.³ In sum, this version of the Ramsey sentence merely says that our theory is empirically adequate in all nomologically possible worlds, plus a cardinal claim.

This formulation does not say how possible worlds should be interpreted metaphysically speaking. At the very least, we are talking about natural necessity, not conceptual or logical necessity, but modal relations could be causal relations, or primitive laws, or relations between universals. Or perhaps nomologically possible worlds are just as concrete as the actual world. In all these cases, possible world semantics can be considered appropriate as a formal tool. Therefore, the conclusion we reach is neutral with regard to the metaphysics of modalities, at least assuming possible world semantics. If we used a possible situation semantics, with the notion of possible situation introduced in Chap. 5, it would differ slightly: the theory would be empirically adequate for all possible situations. However, we can ignore this difference for now.

At this point, the question is the following: should a structural realist be satisfied with such an account? Is it enough to get “real” relations? The answer to this question depends on what criteria we use: what do we mean by “real” relations?

On a first approach, we could stick to the arguments that structural realism purports to answer. We can claim that this solution does the job if (i) modal relations survive theory change and (ii) they are sufficient to *explain* the empirical success of theories. Point (ii) is crucial to know if modal structural realism is distinct from modal empiricism, since one of the characteristic differences between realism and empiricism, as we have seen in Chap. 6, is that the latter does not pretend to explain empirical success while realism, including structural realism in its different guises, does (see for example Ladyman and Ross 2007, p. 79). Let us examine this. Then we will return to point (i).

7.3.2 *Are Modal Relations Real?*

In what sense would the modal Ramsey sentence (or its transposition to the semantic view) constitute an explanation for the empirical success of theories? According to scientific realism, the entities postulated by a theory explain the observable phenomena the theory accounts for. But here, rather than positing real entities, it seems that we have merely extended empirical adequacy to other possible worlds, and it is dubious that such an extension could constitute an appropriate explanation (see Sect. 5.3.2).

Rather than focusing on possible worlds, one could claim that we are postulating nomological constraints on the phenomena, and that these constraints are the entities that explain observable regularities in the actual world. But in what sense are such nomological relations real? Reality is generally cast in terms of mind-independence:

³I will not engage here in the debate between possibilism and actualism, that is, if we should quantify over the same objects or different objects in all possible worlds.

what is real is what does not depend on the way it is represented. We have two options here: either these relations are real because they are relations between real, mind-independent entities, or they are real “in themselves”, that is, the relations themselves are mind-independent.

The first option corresponds to how epistemic versions of structural realism, following the lead of Poincaré and Russell, traditionally conceive of real relations. We saw that Newman’s objection undermines this view: inaccessible objects play no role in structural realism, since they only have to be in sufficient number to bear the structure of the theory, so the relevant relations to which structural realism is committed are actually relations between our observations only. As we saw, the same argument applies to modal structural realism as well: the relations to which modal structural realism is committed are only relations between our observations, and inaccessible objects play no particular role.

Perhaps, though, possible observations could count as real entities? One could argue that possible observations are mind-independent, because they are never actually observed. If such was the case, modal relations would qualify as relations between real objects. But note that past, future or remote observable phenomena are also mind-independent in this sense. They do not have to be actually observed to exist. However, an empiricist like van Fraassen would accept that our theories are empirically adequate for all observable phenomena in the universe, including phenomena that are not actually observed.

The crux here is that mind-independence can be understood in two ways: something is mind-independent either if it does not depend on actually being observed or represented, or if its nature is independent of our conceptualisation of it (but could still correspond to it). We can get the first kind of mind-independence with possible observations, but not the second kind, since possible observations are still conceptualized as observations, and if constructive empiricism is not a brand of realism, the first kind of mind-independence is insufficient for genuine realism.

Now, perhaps van Fraassen is a realist about all observable phenomena, including actually observed ones: perhaps he assumes that their nature is independent of our conceptualisation of it. This would mean that constructive empiricism is a version of structural realism as well. However, this does not look like a helpful way of putting things, given the divergence between these respective positions and the argumentative strategies employed to defend them. It makes more sense, given how the debate is framed, to reserve the term “realism” for the kind of postulated entities that are not directly accessible by their very nature, but that purport to *explain* our observations.⁴

If modal relations are not real in virtue of their relata, this leaves us with the second option, which is to assume that they are real *in themselves*. In other words, the solution is to move to an ontic version of structural realism, following the lead

⁴In any case, as explained in Sect. 4.4.3, modal empiricism is not committed to the notion of observable phenomena entertained by van Fraassen, so it is not realist about observable phenomena, not even in this sense.

of Cassirer (1937) and Eddington (1955) rather than Poincaré and Russell, whereby relations are primitive, and either the relata are fully determined by the relations in which they take part, or the two are in a co-determination relationship (see French 2014 for a review). Cassirer and Eddington conceived of relations as mental entities, but a realist can conceive of these relations as primitive ontological entities, and assume a correspondence between the structure of our representations and reality. However, there are problems with this move.

A problem that has been discussed at length in the literature is that it is hard to make sense of relations without relata, and still make a distinction between mathematical and physical structure (see for example Cao 2003; Psillos 2015). Defenders of ontic structural realism want modalities to be the distinctive feature, and at the same time, they assume that the relata of the structure are fully determined by their position in the structure. But how shall we make sense of this? Take the example of a probabilistic law that would say that if A is the case, then B will occur with probability p and B' will occur with probability $1 - p$, where A , B and B' are observable states of affairs. If following ontic structural realism the modal structure is “all there is”, then something is obviously missing in our ontology: the fact that in such a situation, B actually occurs rather than B' is not entailed by the law itself. This objection applies notably to an interpretation of probabilistic theories, such as quantum mechanics. A natural response consists in reducing this aspect to an indexical component by being a modal realist in the sense of Lewis (1973) (here restricted to nomologically possible worlds): B and B' both occur in different concrete possible worlds, and the fact that B actually occurs only means that we are located in a world where B occurs. In the context of quantum mechanics, this means adopting a many-world interpretation (with potential problems associated with the interpretation of probabilities). But with this solution, it seems that the law is no longer primitive: it supervenes on real states of affairs distributed in different possible worlds. Or if the nature and distribution of these states of affairs are fully determined by the laws themselves, it's not clear that structural realism does not collapse into some form of Pythagoreanism: all we are left with is a mathematical structure of concrete, possible worlds.

In light of these remarks, it seems that we must assume that actual relata exist on a par with a primitive modal structure for this structure to be “modal” in a meaningful sense, and not a pure mathematical abstraction. These actual relata must not be entirely determined by the structure. Psillos (2015) makes a similar point, noting that laws of nature do not determine the initial conditions of the universe. According to Cao (2003), a structure must concern “physical inputs” described qualitatively in order to count as physical.

In response, French (2014, p. 105) implies that this discussion stems from a misunderstanding: “the structure that the structural realist is concerned with should not be, and never should have been, construed as ‘pure’ logico-mathematical structure”. The distinction between mathematical and physical structure could be understood in terms of manifestation: a physical structure is manifested by “existential witnesses” that exist on a par with the structure (this draws on the distinction between determinable and determinate) (French 2014, ch. 8.2, 8.5, 10.7).

The inclusion of these existential witnesses into our ontology gives us a sense of what being a modal structure amounts to. But is it enough to distinguish between mathematical and physical structure?

The problem is reminiscent of Newman’s objection. In order to count as real, a relation, whether it is modal or not, needs to be interpreted. However, if it is only interpreted empirically, if, for example, existential witnesses are potential observations, then we are back to square one: the structure is qualified in mind-dependent terms, and what we have is a modal empiricism. Ontic structural realism, in French’s eliminativist version, does not take existential witnesses to be real, inaccessible objects, and rightly so: such a position would not be informative, for the reasons given by Newman’s objection. At this point, the only remaining option in order to maintain a realist position and eschew Pythagoreanism seems to qualify the structure itself, and in a sense, being “manifested” and “primitively modal” could be viewed as mere qualifications. The risk, as we saw in Sect. 7.2.2, is for this qualification to be either too strong to survive theory change, or too weak to escape triviality.

This leads us back to point (i) above. In the following, I wish to show that if relations are qualified, so as to be considered real “in themselves”, then they cannot survive theory change.

7.3.3 *Which Modal Relations Are Retained in Theory Change?*

There are intuitive reasons to believe that (i) is true, that is, that modal relations survive theory change. If anything is retained in theory change, it is what Duhem (1906) called “experimental laws”. Theories of light have changed several times since the seventeenth century, but they all had to account for light reflection and refraction: they have just embraced more and more types of phenomena with time (diffraction, interference, ...). Now, assuming that such experimental laws express constraints of necessity in the phenomena they describe apparently does nothing to alter this continuity. At most, new theories will restrict the domain to which the old theory is applicable, so that the constraints of necessity between observable phenomena expressed are only approximate (for example, old theories of light neglect magnetic influences, and are only valid when there is no magnetic field). But if the new, wider-range constraints are considered physically necessary, then one could think that the old, narrower constraints are physically necessary too, at least restricted to their domain of application.

All this seems intuitively true, and I have defended similar claims in the previous chapters, assuming the notion of relative necessity defined in Sect. 5.2.2. However, it does not go without saying if one assumes a notion of necessity that is strong enough to qualify as “real”, for example, nomological necessity.

Let us start from the premise that a new, wide-range constraint on phenomena is a matter of nomological necessity, and that it entails some constraint between phenomena *A* and *B* in the narrow context where background conditions *C* are

present. A could be the fact that there is an incident ray, B the fact that there is a reflected ray with the same angle, and C the fact that there is no magnetic field. This can be expressed as follows, with modal operators indicating nomological necessity:

$$\Box(C \rightarrow (A \rightarrow B))$$

However, it is not possible, from this premise, to deduce the following:

$$C \rightarrow \Box(A \rightarrow B)$$

This could only be deduced if C was necessary (in which case $\Box(A \rightarrow B)$ could also be deduced!), but unless all facts are necessary, which would trivialize our position, we cannot assume it: the fact that there is no magnetic field at some place in the universe is not necessary, but contingent.

In other words, if the new theory says: “It is necessary that with background conditions C , there is a relation between A and B ”, that does not entail that with background conditions C , there is a necessary relation between A and B . This is not so much of a problem for a modal empiricist, who could say that the old theory is still modally adequate if we pragmatically restrict the range of possibilities to the ones where C is the case. This is permitted once one adopts a possible situation semantics instead of a possible world semantics: the statement can be made true by restricting the range of situations considered to the ones where C holds. C is considered necessary in context, in virtue of the way relevant situations are identified (so, the notion of necessity involved for fixing C is not metaphysical, but conventional: see Sect. 5.2.1). But for a structural realist who would like the necessity involved to be nomological, the fact that there is no necessary relation between A and B , but only between A , B and C , is problematic because the relation between A and B does not correspond in any sense to the fundamental, modal structure of the world: it only corresponds to the structure of phenomena in situations where C holds. In other words, it is relative to a context where C is the case. But the fact that C holds in some places, and not others, is contingent, not necessary.

Perhaps the law of the old theory could be seen as an approximation of the new law. Let us say that a modal statement approximates a law of necessity if it is true in “most possible contexts”, where the context refers here to background conditions. The problem with this strategy is that it requires one to provide a measure on contexts, and it is not obvious how to do so, and whether it makes sense at all. Different background conditions could correspond to different models in the new theory, and a theory does not provide a measure on its models. Even granting that it makes sense, it does not seem that the law of an old theory of light that, say, does not take into account magnetic influences, would approximate the law of the new theory, except if we chose an ad-hoc measure on contexts. Arguably, there are many more possible configurations of the magnetic field where the old law would fail than configurations where the old law would succeed. In light of theory change, the success of the old law is better accounted for by the fact that we live in places where

the magnetic field is weak than by the fact that it could approximate a strict law of necessity. So the relations that are retained in theory change are either contingent, or if modal, they are so only in a relative sense, such as the sense adopted by modal empiricism, but they are not absolutely modal, nor do they approximate absolutely modal relations.

Maybe this will be clearer with an example. Take Galileo's law of free fall, according to which all falling objects accelerate at 9.8 m/s^2 . The law is certainly retained in contemporary physical theories in the sense given by Post's principle of correspondence mentioned above. However, there is no sense in which it is a modal relation pertaining to the fundamental structure of the world: the number "9.8", although it is part of the mathematical structure of this law, does not appear in contemporary physics, and not even the idea of constant acceleration is retained. Arguably, this law only approximates contemporary ones in a tiny range of possible contexts with specific background conditions. This law is rather, in light of new theories, a contingent empirical consequence of the fact that we live on the surface of the Earth, and that the Earth has a given mass and a given size, but all these contextual aspects, and the way they are involved, are only accessible through the new theories.

There is no problem, from an empiricist perspective, in making this kind of relation relative to our epistemic situation in the world, since, after all, an empiricist is often willing to accept that the content of our representation is not absolute but relative to our epistemic position. But this epistemic relativity is not compatible with a realist stance towards these relations of necessity. This is where the main difference between modal empiricism and structural realism lies. Postulating a relation of correspondence between Galileo's law and the modal structure of the world is problematic given that it appears to be merely contingent. It is not true that modal relations are retained in theory change, or at least, not the kind of modal relation a realist would call for. Either the relations that survive theory change are not modal, but contingent, or if they are modal in a restricted sense, they are not primitive nor nomological, but relative to our specific epistemic position. The structural realist can accept that the laws of past theories, such as Galileo's law of free fall, are not fundamental, of course, but then, by meta-induction on past theories, there is no reason to think that the laws of contemporary theories are fundamental either.

French (2020) addresses this argument, and provides a fallibilist answer. I will examine it later, together with other responses to the pessimistic meta-induction.

7.4 Relativity and Fundamentality

I have argued in the previous section that modal relations only survive theory change if one adopts a notion of relative necessity, and that it is incompatible with realism. This analysis applies quite straightforwardly to versions of structural realism that maintain that our physical theories represent the fundamental modal structure of

the world, given that a relatively modal relation $\Box R$ cannot count as fundamental in a world where $\Box(C \rightarrow R)$ is the case. This applies in particular to French's version and its repeated emphasis on fundamentality (French 2014, p. 44, p. 114, ch. 10). However, structural realism, even in its ontic version, is not a monolithic doctrine, and one could wonder whether less metaphysically inclined versions could accommodate relative modalities.⁵

7.4.1 *Real Patterns and Locators*

According to Ladyman and Ross (2007), the modal relations science describes need not be fundamental. In particular, the “real patterns” discovered by special sciences are specific to certain domains or scales and are relative to (“projectible under”) perspectives (Ladyman and Ross 2007, ch. 4.4). Ladyman and Ross go as far as claiming that they can make sense of relative necessity (Ladyman and Ross 2007, p. 288). This seems at odds with our contention that modal relations that are relative to our epistemic context cannot count as real.

Note that the relativity invoked by Ladyman and Ross is not enrolled in discussions on theory change, but rather on scales and fundamentality and the relationship between physics and the special sciences. In their account, the perspective to which modal relations are relative can be specified by “locators”. However, the problem we are confronted with is not so much a problem of fundamentality. The notion of locators Ladyman and Ross employ seems inappropriate for our purpose. This notion is apparently extensional (Ladyman and Ross 2007, ch. 2.3.4). It amounts to limiting the domain of a theory to part of the actual universe. The parthood relation involved need not be spatio-temporal: it could be, for example, “at a certain scale”, or “for a particular set of actual phenomena”. However, old and new theories of light both apply to the same type of phenomena, at the same scale, and as we have seen in the previous section, limiting the domain of a theory to parts of the actual universe where C is the case will not make its modal statements true, unless C is necessarily the case in those parts of the universe. For example, limiting the domain of an old theory of light to parts of the universe where there is no magnetic field will not make its laws true as a matter of nomological necessity, in so far as there could have been a strong magnetic field in those places. Or limiting the domain to situations near the surface of the Earth does not make Galileo's law necessarily true, because the Earth could lose part of its mass following a cosmic disaster. What needs to be fixed, if one wants to have relations that survive theory change, is not the location, but the background conditions.

To say it differently, the appropriate limitation for our modal statements is not extensional, but intensional: it is a limitation in terms of possible contexts, where some background conditions that are not specified by the theory must hold

⁵I am thankful to anonymous referees for pressing me on this issue.

in all relevant possibilities. Invoking locators might be fine for the prospect of understanding the role of special sciences with respect to more fundamental theories (although I suspect that intensional constraints are involved as well), but in the context of theory change, the relevant intensional context is only accessible in light of a new theory, and so, we are in no position to claim that contemporary theories are structurally true relative to specific background conditions. All we can do is allude to our epistemic position.

One could argue that locators “point” to the relevant background conditions, by selecting situations of interest and assuming a *ceteris paribus* clause. I have proposed something along this line in Sect. 5.2.1. The problem is that pointing to background conditions does not make them explicit, and an ineliminable epistemic relativity remains. Indeed, claiming that the relations that our theories describe are “relatively modal” without specifying the background conditions to which they are relative would lead to triviality in the context of Newman’s objection, since quite any relation can be considered modal relative to some (possibly contrived) conditions. This shows that relatively modal relations understood in this way cannot count as real.

Schurz (2009) argues against the pessimistic meta-induction by proving a theorem according to which under some conditions of empirical success, old theories refer “indirectly” to entities posited by new theories. This is taken to support a form of structural continuity between theories. This notion of indirect reference, and the associated notion of partial truth that Schurz employs, incorporate a domain-relativity of the kind examined here: the idea is that the old laws can be expressed as the restriction of the new laws to a particular domain of experience. It is thought that ultimately, our theories indirectly refer to the entities posited by a true theory. However, this notion of indirect reference is insufficient for realism for the reasons just given: relative laws cannot count as real. Indeed, by the standards set up by Schurz, the structure of Galileo’s law of free fall would count as an indirect reference to the structure of general relativity. It might be true that all the laws of contemporary physics are expressible as the restriction of the laws of a true theory to a particular domain. However, since this domain remains unspecified by our current theories, this is not informative.

In sum, the patterns discovered by special sciences could count as real if the background conditions to which they are relative could be specified, presumably in a more fundamental theory. In other words, we still need a more fundamental level that is not relative to our epistemic position if we want to make sense of Ladyman and Ross’s position. Ladyman and Ross are consistent in this respect: even though they assume that special sciences describe real patterns that are only present at certain scales, they entertain a more ambitious stance towards physics. According to them, fundamental physics discovers structures of a “higher level of necessity”. It is responsible for supporting counterfactuals across the entire actual universe: the residual relativity of modal statements is a relativity to contingent “structural facts about the whole universe”, such as its initial conditions. This is why fundamental physics “gives the modal structure of the world” (Ladyman and Ross 2007, p. 288). Such a stance towards physics seems necessary for structural realism to qualify as a

version of realism. The problem is that the arguments of the previous section apply to this view: this modal structure does not survive theory change.

7.4.2 *Fallibilism*

Ladyman and Ross (2007, p. 99)'s claim that "ontic structural realism ought to be understood as modal structural empiricism" is instructive. Why classify this position as a brand of realism? For them, "[modal empiricism] is a form of structural realism because according to it the theoretical structure of scientific theories represents the modal structure of reality" (Ladyman and Ross 2007, p. 111) (I will take the faithfulness of this representation to be implicit here). What a faithful representation is, and how it latches onto reality, can be understood in different ways in a structuralist context (see French 2014, ch. 5.10), but one would expect that a notion of truth (perhaps partial or approximate truth) that is not epistemically constrained, such as correspondence truth, would be involved at some point. Arguably, an epistemically constrained notion of truth, such as pragmatic truth, is incompatible with realism as generally understood.⁶ For example, Ladyman and Ross take a pragmatic stance towards individual objects, the existence of which they deny.

This adoption of a strong notion of truth should hold even within a fallibilist approach, wherein structural realists would grant that theoretical modal statements are revisable and that contemporary theories of fundamental physics are not at the stage of composing a "grand unified theory" (as exemplified by French 2014, p. 163) (French 2020).

As observed in Sect. 6.4.2, this fallibilist approach can be adopted by an empiricist, because in light of new theories, one can easily claim that old theories are still empirically adequate within a limited domain of application (in a range of possible situations). Past theories are still modally adequate in this domain and science advances by accumulating empirical knowledge. Scientific progress can be viewed as an extension of the range of epistemic contexts available to us, which does not mean that we could be free of any epistemic context at any point in time, or that we could claim to have access to "most possible epistemic contexts", whatever that means.

However, fallibilism is much more problematic for a structural realist. I assume that this stance makes sense if the fallibilist assumes that contemporary theories still approximately capture part of the modal structure of reality, but as we have seen previously, there is no sense in which a statement $\Box R$ could strictly, partly or approximately correspond to a modal structure described by $\Box(C \rightarrow R)$, so it does

⁶Some positions that adopt an epistemically constrained notion of truth are labelled realism. This is the case, for example, of Putnam's internal realism. Modal empiricism could qualify as realism if the term is understood in this very broad sense, but this is not the general acceptance of the term. This is a terminological issue that will be addressed in Chap. 8.

not make much sense to claim that old theories are still true in a limited domain. By meta-induction on past theories, we have every reason to assume that current theories are as false as past ones. An honest fallibilist should therefore consider that it is likely that our theories are not even approximately true (whereas they can be approximately adequate), but then it is not clear what realist component remains.

I should make a remark on the idea briefly mentioned above that we would have access to “most possible epistemic contexts”. In response to the pessimistic meta-inductions, various authors have argued that science has grown exponentially, that we are now much better at making observations than before, that we have explored more alternative theories, and so, that contemporary theories are much more likely to be true than the ones of the past (Ruhmkorff 2013; Fahrbach 2011; Devitt 2011; Doppelt 2007). The main problem with this argument is that we lack an external viewpoint in order to evaluate this idea. A few centuries ago, the universe was thought to be much smaller than it is thought to be now, and scientists of the past could entertain the idea that they had explored most of it. I see no reason to assume that we are in a different position now, and I do not think that we have any means of knowing that. This is a good reason for suspending our judgment in these matters (see also Wray 2013; Müller 2015).

In light of this, we should assume, at most, that our theories capture the modal structure of observations that are accessible from our epistemic context, which corresponds to modal empiricism. This ineliminable reference to an epistemic context associated with unknown background conditions that we are in no position to specify any further until a new theory arrives is precisely what makes modal empiricism a version of empiricism rather than realism: only an epistemically constrained notion of truth, such as pragmatic truth, will account for it.⁷

In sum, Ladyman and Ross’s version of ontic structural realism might be as close to modal empiricism as structural realism can get, but their contention that physics gives us the modal structure of the world makes it a distinct position, and unless they provide an account of partial or approximate truth for relative modalities that solves the problems mentioned here, their position cannot be sustained in the face of the pessimistic meta-induction.

7.5 Modal Empiricism: The Best of Both Worlds

The aim of this chapter was to highlight the main difference between modal empiricism and scientific realism. This was done by comparing modal empiricism to the closest version of realism: structural realism. As we can see, the difference

⁷There is a further and more radical difference between modal empiricism and structural realism that was not addressed in this chapter. The normative component of communal representations, associated with the performative component of contextual representation, could imply that scientific models and theories are not purely descriptive. I will develop this theme in the next chapter.

between the two lies in the interpretation of modalities. Modal empiricism is committed to the existence of situated possibilities, in the sense developed in Chap. 5. It entails that the relations of necessity that are discovered are only relative to the situations of the world to which we have access, and to the aspects of these situations that are accessible to us. In sum, according to modal empiricism, (1) relations of necessity are not absolute, but relative to an unspecified background context associated with our epistemic position, and (2) they are relations between possible observations and manipulations, which are also relative to our cognitive abilities and (in contexts) particular purposes. Since neither the relations nor the relata to which modal empiricism is committed can be considered real, modal empiricism is not realism. It does not purport to explain the empirical success of theories by appealing to real relations: it merely states this success.

This makes a crucial difference when addressing the pessimistic meta-induction. An empiricist stance is compatible with a fallibilist approach towards theories, because failures are easily accommodated by restricting the domain of application of our theories. I am not convinced that fallibilism really makes sense for a realist if failure means that theories are false simpliciter, and not approximately or partly true. As we have seen, a commitment to “real” relations, be it nomological or “primitively modal” relations, has this consequence.

Worrall claimed that structural realism offers us the “best of both worlds” between scientific realism and empiricism, because it can respond to the two main arguments in the debate: the pessimistic meta-induction and the no-miracle argument. We have seen in the previous chapter how modal empiricism could respond to the no-miracle argument, and in this chapter, we have seen that it does not fall prey to the pessimistic meta-induction in the same way realism does. I believe that structural realists are right that one can have the best of both worlds by focusing on the modal structure of theories, and I would say that it is what modal empiricism achieves. The only mistake of structural realists is to position themselves in the realist camp, while if they want to address problems of theory change, they simply have to be empiricists.

The notion of truth and its various conceptions came up during the discussions of this chapter. The difference between realism and modal empiricism could indeed be addressed in terms of truth: what makes a statement of necessity true, according to a modal empiricist, is not a correspondence to a nomological structure, because it is dependent on a perspective. Perhaps then modal empiricism could be expressed differently. Instead of claiming that theories are not true, but merely empirically adequate (or that the aim of science is to produce such theories), we could say that theories are true, but in a pragmatist sense. I will explore this proposal in the next and final chapter of this book.

References

- Ainsworth, P. (2009). Newman's Objection. *British Journal for the Philosophy of Science*, 60(1), 135–171.
- Bueno, O. (2008). Structural realism, scientific change, and partial structures. *Studia Logica*, 89(2), 213–235.
- Cao, T. Y. (2003). Can we dissolve physical entities into mathematical structures? *Synthese*, 136(1), 57–71.
- Cartwright, N. (1983). *How the laws of physics lie* (vol. 34). Oxford: Oxford University Press.
- Cassirer, E. (1937). *Determinismus Und Indeterminismus in der Modernen Physik Historische Und Systematische Studien Zum Kausalproblem*. Mölnlycke: Elanders Boktryckeri Aktiebolag.
- Cei, A., & French, S. (2006). Looking for structure in all the wrong places: Ramsey sentences, multiple realisability, and structure. *Studies in History and Philosophy of Science Part A*, 37(4), 633–655. <https://doi.org/10.1016/j.shpsa.2006.09.006>
- Chang, H., & Leonelli, S. (2005). Infrared metaphysics: The elusive ontology of radiation. Part 1. *Studies in History and Philosophy of Science Part A*, 36(3), 477–508.
- Clarke, S. (2001). Defensible territory for entity realism. *The British Journal for the Philosophy of Science*, 52(4), 701–722. <https://doi.org/10.1093/bjps/52.4.701>
- Cruse, P. (2005). Ramsey sentences, structural realism and trivial realization. *Studies in History and Philosophy of Science Part A*, 36(3), 557–576.
- Demopoulos, W., & Friedman, M. (1985). Bertrand Russell's the analysis of matter: Its historical context and contemporary interest. *Philosophy of Science*, 52(4), 621–639.
- Devitt, M. (2011). Are unconceived alternatives a problem for scientific realism? *Journal for General Philosophy of Science/Zeitschrift für Allgemeine Wissenschaftstheorie*, 42(2), 285–293.
- Doppelt, G. (2007). Reconstructing scientific realism to rebut the pessimistic meta-induction. *Philosophy of Science*, 74(1), 96–118.
- Duhem, P. (1906). *La théorie physique: son objet, et sa structure*. Chevalier & Rivière.
- Eddington, A. (1955). *The nature of the physical world* (vol. 39). London: Dent.
- Fahrbach, L. (2011). How the growth of science ends theory change. *Synthese*, 180(2), 139–155. <https://doi.org/10.1007/s11229-009-9602-0>
- French, S. (2014). *The structure of the world: Metaphysics and representation*. Oxford: Oxford University Press.
- French, S. (2020). *What is this thing called structure? (Rummaging in the Toolbox of Metaphysics for an Answer)*. <http://philsci-archive.pitt.edu/16921/>
- French, S., & Ladyman, J. (2003). Remodelling structural realism: Quantum physics and the metaphysics of structure. *Synthese*, 136(1), 31–56.
- Frigg, R., & Votsis, I. (2011). Everything you always wanted to know about structural realism but were afraid to ask. *European Journal for Philosophy of Science*, 1(2), 227–276.
- Greaves, H., & Wallace, D. (2014). Empirical consequences of symmetries. *The British Journal for the Philosophy of Science*, 65(1), 59–89. <https://doi.org/10.1093/bjps/axt005>
- Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science* (vol. 22). Cambridge: Cambridge University Press.
- Kosso, P. (2000). The empirical status of symmetries in physics. *The British Journal for the Philosophy of Science*, 51(1), 81–98. <https://doi.org/10.1093/bjps/51.1.81>
- Ladyman, J. (2004). Constructive empiricism and modal metaphysics: A reply to Monton and Van Fraassen. *British Journal for the Philosophy of Science*, 55(4), 755–765.
- Ladyman, J., & Ross, D. (2007). *Every thing must go: Metaphysics naturalized*. Oxford: Oxford University Press.
- Laudan, L. (1981). A confutation of convergent realism. *Philosophy of Science*, 48(1), 19–49.
- Lewis, D. (1973). *Counterfactuals*. Hoboken, NJ, USA: Blackwell Publishers.
- Lyons, T. D. (2017). Epistemic selectivity, historical threats, and the non-epistemic tenets of scientific realism. *Synthese*, 194(9), 3203–3219. <https://doi.org/10.1007/s11229-016-1103-3>

- Maxwell, G. (1970). Theories, perception and structural realism. In R. Colodny (Ed.), *The Nature and Function of Scientific Theories* (pp. 3–34). Pittsburgh: University of Pittsburgh Press.
- Maxwell, G. (1971). Structural realism and the meaning of theoretical terms. *Minnesota Studies in the Philosophy of Science*, 4, 181–192.
- McGuire, J. (1970). Atoms and the 'Analogy of Nature': Newton's third rule of philosophizing. *Studies in History and Philosophy of Science Part A*, 1(1), 3–58.
- Melia, J., & Saatsi, J. (2006). Ramseyfication and theoretical content. *British Journal for the Philosophy of Science*, 57(3), 561–585.
- Morrison, M. (1990). Theory, intervention and realism. *Synthese*, 82(1), 1–22. <https://doi.org/10.1007/BF00413667>
- Müller, F. (2015). The pessimistic meta-induction: Obsolete through scientific progress? *International Studies in the Philosophy of Science*, 29(4), 393–412.
- Newman, M. (1928). Mr. Russell's causal theory of perception. *Mind*, 5(146), 26–43.
- Newman, M. (2005). Ramsey sentence realism as an answer to the pessimistic meta-induction. *Philosophy of Science*, 72(5), 1373–1384.
- Papineau, D. (1996). *The philosophy of science*. Oxford: Oxford University Press.
- Pashby, T. (2012). Dirac's prediction of the positron: A case study for the current realism debate. *Perspectives on Science*, 20(4), 440–475.
- Poincaré, H. (1902). *La science et l'hypothèse*. Bibliothèque de philosophie scientifique, E. Flammarion.
- Post, H. (1971). Correspondence, invariance and heuristics: In praise of conservative induction. *Studies in History and Philosophy of Science Part A*, 2(3), 213–255.
- Psillos, S. (1995). Is structural realism the best of both worlds? *Dialectica*, 49(1), 15–46.
- Psillos, S. (1999). *Scientific realism: How science tracks truth*. Philosophical Issues in Science. London: Routledge
- Psillos, S. (2001). Is structural realism possible? *Proceedings of the Philosophy of Science Association*, 2001(3), 13–24.
- Psillos, S. (2015). Broken structuralism: Steven French: The structure of the world: Metaphysics and representation. *Metascience*, 25, 163–171. <https://doi.org/10.1007/s11016-015-0030-0>
- Redhead, M. (2001). The intelligibility of the universe. *Royal Institute of Philosophy Supplement*, 48, 73–90.
- Ruhmkorff, S. (2013). Global and local pessimistic meta-inductions. *International Studies in the Philosophy of Science*, 27(4), 409–428.
- Russell, B. (1927). *The analysis of matter*. London: Kegan Paul.
- Ruyant, Q. (2019). Structural realism or modal empiricism? *The British Journal for the Philosophy of Science*, 70(4), 1051–1072. <https://doi.org/10.1093/bjps/axy025>
- Ruyant, Q. (2020). Symmetries, Indexicality and the Pragmatic Stance
- Saatsi, J. (2017). Replacing recipe realism. *Synthese*, 194(9), 3233–3244. <https://doi.org/10.1007/s11229-015-0962-3>
- Saatsi, J., & Vickers, P. (2011). Miraculous success? Inconsistency and Untruth in Kirchhoff's diffraction theory. *The British Journal for the Philosophy of Science*, 62(1), 29–46. <https://doi.org/10.1093/bjps/axq008>. <https://academic.oup.com/bjps/article-lookup/doi/10.1093/bjps/axq008>
- Schlick, M. (1932). *Forme & contenu: une introduction à la pensée philosophique*. Agone éditeur.
- Schurz, G. (2009). When empirical success implies theoretical reference: A structural correspondence theorem. *British Journal for the Philosophy of Science*, 60(1), 101–133.
- Stanford, K. (2003). Pyrrhic victories for scientific realism. *Journal of Philosophy*, 100(11), 553–572.
- Stanford, P. K. (2010). *Exceeding our grasp: Science, history, and the problem of unconceived alternatives*. New York; Oxford: Oxford University Press.
- Thébault, K. P. Y. (2016). Quantization as a guide to ontic structure. *The British Journal for the Philosophy of Science*, 67(1), 89–114. <https://doi.org/10.1093/bjps/axu023>
- van Fraassen, B. (2002). *The Empirical stance*. The Terry Lectures. New Haven, CT, USA: Yale University Press.

- Votsis, I. (2004). *The epistemological status of scientific theories: An investigation of the structural realist account*. Doctoral dissertation, London School of Economics and Political Science (United Kingdom).
- Worrall, J. (1989). Structural realism: The best of both worlds? *Dialectica*, 43(1–2), 99–124.
- Wray, K. B. (2013). The pessimistic induction and the exponential growth of science reassessed. *Synthese*, 190(18), 4321–4330.
- Zahar, E. (2004). Ramseyfication and structural realism. *Theoria*, 19(1), 5–30.

Chapter 8

Semantic Pragmatism



Abstract Modal empiricism can simultaneously provide a faithful picture of the functioning of science and respond to the epistemic challenges of the debate on scientific realism. If we accept it, why maintain a gap between our semantic theories concerning the content of scientific theories and what we take to be the achievements of science? Why not equate theoretical truth with modal empirical adequacy, and theoretical meaning with conditions of ideal modal empirical success? This concluding chapter examines the issue of semantic realism, and suggests that an alternative pragmatist conception, based on modal empirical adequacy, could well account for the semantics of scientific discourse. As a consequence, a metaphysician should not shy away from being a modal empiricist if she accepts reinterpreting her activity. The main advantage of this proposal is that it is able to bring pragmatic relevance to metaphysical discourse.

8.1 Anti-Realism, Acceptance and Belief

In this book, I have presented and defended modal empiricism, which is a position of compromise in the debate on scientific realism. According to modal empiricism, there are possibilities in the world, and natural constraints on these possibilities. The aim of science is to produce theories that correctly account, in a unified way, for the way these constraints affect our possible observations and actions. Science is generally successful in this aim, and scientific theories thus allow us to navigate successfully in this world: this is the kind of understanding that science has to offer.

Modal empiricism puts the contextual aspects of scientific practice at centre stage, and understands communal aspects in relation to the former rather than the other way around. Scientific theories and theoretical models are understood as conveying communal norms for their possible contextual applications, so the aim of science is to give us ideally successful norms of representation. This stance towards theories is motivated by recent debates on scientific representation, which were presented in Chap. 2, and it is articulated in the form of an account of epistemic representation, which was developed in Chap. 3.

In the course of this book, I have defended modal empiricism against its main opponents, constructive empiricism, scientific realism and structural realism, by examining the arguments of the axiological and epistemic debates on scientific realism. In Chap. 4, I developed the notion of empirical adequacy that characterises modal empiricism and compared it to the one proposed by van Fraassen, in order to argue that modal empiricism is better apt to make sense of scientific practice as a rational activity. This is the axiological component of the debate, the one concerned with the aim of science. The main difference between constructive empiricism and modal empiricism is that modal empiricism expresses empirical adequacy in terms of situated uses, and in modal terms: an ideally successful theory should account for merely possible observations and manipulations in particular situations. An adequate theory gives us a reliable map of what it is or is not possible to do and observe in this world in what circumstances, from our perspective. This helps account for the interventionist component of scientific experimentation. In Chap. 5, I presented in more detail the situated modalities to which modal empiricism is committed. I gave a few semantic, pragmatic and empirical reasons to accept them, and rebutted sceptical arguments against the possibility of modal knowledge by proposing an inductivist epistemology for situated modalities. This result supports the soundness of taking modal empirical adequacy to be the aim of science, and constitutes a first step in the epistemic debate, concerned with the achievements of science and the justification of our attitudes towards scientific theories. It warrants believing in the modal empirical adequacy of successful theoretical models.

In Chap. 6, I examined critically the realist argumentative strategy, and I argued that modal empiricism can respond to the no-miracle arguments and account for successful novel predictions in so far as empirical adequacy can be justified, since novel predictions are among the possibilities for which the theory is adequate. I provided an inductivist justification for empirical adequacy at the theory level, based on an induction on models. Invoking this form of induction threatens to bring modal empiricism very close to scientific realism, since it seems to presuppose that nature is not only uniform, but structured. In Chap. 7, I compared modal empiricism to the closest version of realism, structural realism, and explained that an important difference remains between the two positions. This difference lies in the epistemic relativity of situated modalities. The relations of necessity that are warranted are not absolute, but relative to our position in the universe, and they are presumably intensionally limited. This makes modal empiricism distinctively empiricist, and better apt to respond to problems of theory change. Modal empiricism thus comes out as the best position of compromise in the debate on scientific realism.

This whole debate on whether science aims at truth and whether theories are true presupposes *semantic* realism. As explained in the introduction of this book (Sect. 1.2), semantic realism assumes that scientific theories should be interpreted “at face value” or “literally”, which is taken to imply a robust conception of truth, such as correspondence truth. The arguments in favour or against epistemic realism only make sense assuming such a notion of truth.

Semantic realism is accepted even by some anti-realists. For instance, van Fraassen accepts that scientific theories should be interpreted literally, even though he does not think that the aim of science is to produce true theories in this sense. One could wonder why one should maintain a gap between semantic and epistemic aspects, as van Fraassen does. According to his constructive empiricism, the cognitive content of a theory consists in descriptions of reality, but at the same time, the aim of science is not to describe reality. So, when a scientist presents a theory, we should interpret the content of what she says in terms of literal descriptions of the world, but without presupposing that she actually aims at describing anything, except observable phenomena. This is rather counterintuitive.

In order to make sense of this gap, van Fraassen introduces a distinction between acceptance and belief. According to van Fraassen (1980),

To accept a theory is to make a commitment, a commitment to the further confrontation of new phenomena within the framework of that theory, a commitment to a research programme, and a wager that all relevant phenomena can be accounted for without giving up that theory. (p. 88)

Acceptance also has a linguistic dimension:

The language we talk has its structure determined by the major theories we accept. That is why, to some extent, adherents of a theory must talk just as if they believed it to be true. (p. 202)

However, accepting a theory does not require believing that it is true: “acceptance of a theory involves as belief only that it is empirically adequate” (p. 12), and this belief can be a matter of degree (p. 9). Such acceptance is associated with an empiricist *stance* towards science. Van Fraassen does not claim that we *should* accept theories in this sense, nor that we should not believe in more than empirical adequacy. A stance, unlike a philosophical thesis, involves a voluntary element, and because of this, stances cannot be evaluated on purely neutral grounds. They should be tolerated as long as they are not irrational. This aspect is part of van Fraassen’s voluntarist epistemology (van Fraassen 1989, ch. 7.5).

I agree with Horwich (1991) and Mitchell (1988) that van Fraassen’s notion of acceptance is not very different from belief. In general epistemology, the difference between acceptance and belief is characterised by the following criteria: acceptance, contrarily to belief, is context-dependent; it is subject to voluntary control; it does not come in degrees, and it need not be shaped by evidence (Buckareff 2010). A certain notion of acceptance that corresponds to this understanding of the term can be found in scientific practice: it is the attitude a scientist adopts when performing an empirical test of a theory without presupposing that the theory will pass the test (Maher 1990). However, this does not require believing that the theory is empirically adequate, so this notion is not the one van Fraassen has in mind. Van Fraassen’s notion of acceptance is not context-dependent, because a research programme is not restricted to particular contexts. It comes in degrees. It must be shaped by evidence since a belief in empirical adequacy is involved. Finally, it is not *entirely* subject to voluntary control if this belief is involved.

There might be a meaningful distinction between van Fraassen's notion of acceptance and belief assuming a correspondence theory of truth. However, the idea that it involves "a wager that all relevant phenomena can be accounted for without giving up that theory" makes it very close, if not identical, to the idea that theories are true in the pragmatist sense. In what follows, I will argue that we should adopt such a conception of truth, thus removing the gap between our semantics and our epistemic assumptions, and the need for a distinction between acceptance (in van Fraassen's sense) and belief.

8.2 Correspondence and Pragmatic Truth

As already explained, semantic realism involves a conception of truth that is not epistemically constrained, like correspondence truth. According to this conception, something is true if it corresponds to reality. Such a "verification-transcendent" notion of truth introduces a principled gap between what makes a statement true and our capacities to know that it is true on the basis of empirical evidence, and this gap can be held responsible for the sceptical arguments examined in this book, such as the underdetermination of theories by experience (see Sect. 6.2.1). So, why maintain such a gap between semantic and epistemic aspects? Why is semantic realism so widely accepted?

One reason to be a semantic realist could stem from the idea that correspondence truth is the most natural and intuitive conception of truth. However, correspondence truth has been criticised on various grounds. I will not present all these criticisms here (see Khlentzos 2016), but generally speaking, invoking a transcendental relation between our representations and reality legitimately brings suspicion. When trying to make sense of this idea of correspondence between representations and reality, the best we can do is bring in more representations: one might, for example, represent an agent's belief as a model, and then represent the target of this agent's belief as another model. Then, one will observe that when the represented belief is true, there is a relation of correspondence between these two models, which could be, for example, an isomorphism. Call this representation of a representation relation a *second-order representation*. By doing this, we are not stepping outside of our representations of the world: this isomorphism is still part of a representation. So, claiming that truth is correspondence is not merely holding this second-order representation: one must also claim that this represented isomorphism somehow *corresponds* to a real relation of correspondence between the agent's belief and reality. But again, we could wonder how to analyse this claim, and our only way to do so consists in invoking a third-order representation, which leads to an infinite regress. This transcendental notion of correspondence seems intractable, and one could wonder whether it makes sense at all.

Admittedly, talking about correspondence is an intuitive way of analysing truth, but we need not abandon it. The lesson of these remarks is that we should resist interpreting this intuitive way of talking in a way that transcends our representational

abilities. A scientific description is considered true if it corresponds to its object, but in order to verify if there is correspondence, what we do is compare two representations, and *not* a representation and reality directly (assuming this idea makes any sense at all). One representation could be a data model, or perhaps a perceptual representation of the target, and the other a theoretical model or a statement. Many times, in particular in ordinary contexts, this kind of comparison is enough to claim that a belief is true and that it constitutes knowledge. Sometimes, we say or presume that a description is true without being in a position to verify it, or we consider that its conformity with data is merely an indicator of truth, but that the description might still be false. This merely means that we consider or imagine a comparison between the description and an ideal representation of the object that we do not actually possess. In any case, there is no reason to assume that there is more to correspondence than this kind of comparison between various representations with different sources. In this regard, the correspondence theory of truth seems to trade on an ambiguity between a mundane and a transcendental notion of correspondence.

Does it mean that we should adopt a coherence conception of truth, according to which the truth of a belief is only a matter of coherence with other beliefs? Should we assume that all there is to say about truth is that our various representations of the world all fit in a coherent conceptual scheme? Coherence truth is often associated with idealism. Understood in a minimalist way, it seems to lead to a problematic form of relativism, according to which what is true depends on what one believes. Many incompatible conceptual schemes could be internally coherent, and we seem to lose sight of the world by adopting such a conception of truth. So, we should say more.

This is where pragmatic truth enters the picture. According to the pragmatist, what is important, in order to understand truth, is to focus on the norms of inquiry. We should have a look at how our representations are used, and on what grounds they are accepted or rejected. In this respect, the coherence of our conceptual scheme certainly matters, but not all representations are on an equal footing. In scientific experimentation, the relation between theoretical models and data models is asymmetric: theories are tested against data, and not the other way around. Representations that are closer to experience generally take (or *should* take) priority. Concrete representations, which are rooted in experience, enjoy a form of authority over more abstract representations.

Of course, as is well known, experimental data is revisable. Just as we can doubt that our senses are reliable under some circumstances, particularly if they contradict prior assumptions, we can put into question the accuracy of experimental data in light of a well-established theory. This is what happened during the OPERA experiment, for example, whose results seemed to contradict relativity theory (Reich 2011). However, when experimental data is put into question, correcting data involves checking the experimental setup and the way data was produced, that is, bringing in more information from experience, and not simply adjusting the data to the theory by fiat. Furthermore, a scientific theory only gains this kind of relative authority on particular data when it is highly confirmed, which means that this authority has been gained from many other observations. Being a pragmatist

does not mean being a foundationalist. Coherence matters. However, when deciding how to restore coherence in the face of a contradiction, experience ultimately takes priority over theory (for a similar view, see Israel-Jost (2015)).

This does not mean that truth should be identified with *actual* empirical success, because of course, we could be wrong even when our theories are successful. According to a pragmatist, truth should rather be identified with *ideal* empirical success. A belief is true if it corresponds to an ideal representation of its object, one that could not be defeated by any possible experience. So, we can make sense of the intuitive idea that truth involves correspondence even without being a semantic realist, if we understand correspondence as relating our actual representations to ideal ones. This alternative to semantic realism will be presented in Sect. 8.4.

8.3 Why Semantic Realism?

I guess that the reasons why semantic realism about scientific theories is often uncritically accepted in contemporary philosophy of science have historical roots. Logical empiricists proposed various reductive analyses of the content of scientific theories in terms of observables that did not stand up to scrutiny (see Sect. 2.2). It is much harder to find positive arguments for semantic realism in the philosophy of science literature than it is to find arguments against the reductive semantics of logical empiricists. However, these criticisms sometimes seem to convey the idea that semantic realism would offer the most natural way of interpreting scientific discourse, because scientists would be implicit semantic realists. I have the impression that this is what is implicitly meant when semantic realism is characterised as the idea that we should interpret theories “literally” or “at face value”.

Here is the kind of rationale that is typically offered in the literature. When a scientist talks about electrons, she seems to be talking about real entities that populate the world, and not about detection instruments or their outcomes. Instruments are normally conceived of as means of detecting electrons, not as determining the nature of electrons, and their outcomes are normally understood as being *caused* by electrons. This is how scientists talk, and this way of talking should be taken “at face value”. So, electrons exist. This is an argument against reductive semantics, not against pragmatic truth. A pragmatist can agree that electrons should not be identified with detection instruments, nor with their outcomes, since there are many possible ways, direct or indirect, of detecting electrons, and they are generally context-dependent. Modal empiricism actually accounts for these aspects quite nicely. I will say more about modal empiricism’s interpretation of theoretical terms in the next section. In any case, the failure of particular reductionist attempts does not entail that truth is verification-transcendent, and maybe philosophers have been too fast in their adoption of semantic realism.

One reaction to the failure of the reductionist semantics entertained by logical empiricists has been to move to a model-based approach towards scientific theories (see Sect. 2.2). The idea that such a move would be naturally associated with a correspondence conception of truth is far from obvious. In a model-based approach, theories are not linguistic entities, so theoretical truth must be fleshed out by providing a relation between the models of the theory and reality. Even assuming that there is a well-defined relation of correspondence between particular models and what they represent in particular contexts, it is not obvious how to understand theoretical truth in general.

Claiming that a theory is true if all its models are, or if some of its models are, only makes sense if models are assigned particular targets of representation. So, a theory, in order to be truth-apt, should come with norms that specify how its models can be applied. One can presume that these norms of application rest on a categorisation of phenomena by the theory. The content of a theory will be ultimately expressed in terms of relations between these categories of phenomena and theoretical posits, so a semantic realist must maintain that the way targets of representation are identified and categorised is not epistemically constrained if truth is not epistemically constrained. This is quite hard to swallow: sound norms must be applicable in principle, and their applicability seems to rest on our epistemic and cognitive abilities, so there are *prima facie* good reasons to assume that scientific categories are based on accessible features of phenomena, and that theoretical truth is epistemically constrained. Otherwise, it would be hard to understand how theories can be used at all.¹ User-centred accounts of scientific representation, with their emphasis on the purposes of epistemic agents in theory application, also plead against this view (see Sect. 2.3.2).

A positive argument for semantic realism can be found in the work of Kripke (1980) and Putnam (1975). Since theoretical terms often survive theory change, their meaning cannot be identified with the way particular theories describe them: whales are still whales, whether we classify them as fishes or as mammals. We should therefore interpret theoretical terms in terms of direct reference to natural properties. This reference would be first secured by direct ostentation to particular instances of these properties, and then transmitted from locutor to locutor. The idea that reference to natural properties would be secured by direct ostentation looks plausible in the case of biological species, but less so in the case of physical properties, such as mass or charge. So, this kind of argument for semantic realism is not necessarily conclusive. One could also argue, following LaPorte (2004), that the meaning of theoretical terms does change with theories, but that there are pragmatic reasons to keep a given term in a new theory, for example because the new meaning only makes the old meaning more specific. LaPorte notes, in favour of his thesis, that there were discussions about whether the term “species” should be retained in evolutionary biology after Darwin proposed his theory of evolution, or if we should

¹This is related to Dummett (1978)’s arguments against semantic realism in philosophy of language. See Ruyant (2020) for a more detailed argument. See also Sect. 2.3.2.

assume instead that species do not exist, given that species were often thought not to evolve before Darwin. The way theoretical terms are operationalised often survives theory change, at least in approximation, which could be another pragmatic reason to retain them, since they roughly play the same functional roles in experimentation independently of theories.

The main problem with semantic realism as it is often presented is that the locution “at face value” is inherently vague, and the very idea that interpreting scientific discourse “at face value” implies a form of semantic realism is far from obvious. I am not convinced that scientists or ordinary speakers are implicit semantic realists. Actually, I am not even convinced that they entertain strong views on the semantics of their utterances, let alone strong metaphysical interpretations of these utterances, as the realist usually does.

There is not always one straightforward realist interpretation of a theoretical framework. Take quantum mechanics, for instance. This theory presents us with abstract mathematical structures with different co-existing formulations (wave mechanics versus matrix mechanics), and realist metaphysicians have proposed various incompatible ontologies in order to make sense of the theory, such as point particles, matter-density fields, events or multi-fields: the so-called “primitive ontologies” (Belot 2012).² A consistent semantic realist should say that these proposals are all different scientific theories, and that quantum theory is not a theory at all, because its transcendental truth conditions are underspecified. Calling these proposals scientific theories is counterintuitive, because scientists do not seem as interested in these ontological questions as metaphysicians are. Nevertheless, this conclusion is endorsed by Maudlin (2018). Coffey (2014) also argues for such a notion of interpretation on the basis of limitations of formal accounts of theory equivalence. But some of Coffey’s conclusions seem at odds with ordinary ways of speaking in science. For example, he believes that assuming two different accounts of laws of nature (say, a Humean and a dispositionalist account) when interpreting a given theoretical framework gives rise to two inequivalent theories (footnote 40). The question is: where will interpretation stop? Should different metaphysical conceptions of natural kinds imply a plethora of different theories in chemistry or biology? Semantic realism seems to lead us astray.

Perhaps there is a meaningful way to draw the line between theoretical content and metaphysical content in the context of semantic realism. However, I believe that in order to know what interpreting a theory “at face value” means, it makes more sense to have a look at how the theory is actually interpreted *in context*, when applied to concrete targets of representation by actual scientists, and I suspect that it would give us the notion of equivalence that scientists generally have in mind.

Admittedly, empirical equivalence is too weak to account for the way scientists distinguish between theories. However, the account of representation developed in this work allows for finer pragmatic distinctions (see also Nguyen 2017). For

²Note that these proposals are independent from the interpretative issues associated with the measurement problem.

example, empirically equivalent theories that “permit” the same possible states for targets of representation in all possible contexts could classify contexts differently, by means of general models with different domains of application, which could make a difference in the way these theories can be extended to new domains.³ To give another example, Bohmian mechanics implies that all measurements should be reinterpreted as position measurements, while the standard formulation of quantum mechanics allows for other properties to be measured. Although these two formulations are generally considered to be empirically equivalent, the contexts in which Bohmian models can be interpreted are not delineated exactly in the same way as the ones where standard quantum mechanical models can be interpreted: sometimes, parts of measuring instruments must be represented in a Bohmian model to yield position predictions. So, it seems possible to develop an intuitive notion of theory equivalence and of “face-value” interpretation that is finer than empirical equivalence without resorting to a transcendental relation between theories and reality.⁴

Other observations cast doubt on the idea that ordinary scientific discourse is implicitly realist. Scientists are presumably competent interpreters of their own theories, but at times you can hear physicists claiming that Newtonian mechanics is true “within its domain of validity”. This relativisation of truth to a domain of experience is not clearly compatible with semantic realism. Finally, the standard textbook formulation of quantum theory explicitly mentions measurement, an epistemically loaded notion, while so-called realist interpretations often attempt to complete this standard formulation with additional structure in order to eliminate this dependence on measurements. Invoking an epistemically loaded notion does not seem to cause much trouble from a scientific point of view, and it is far from clear who, in these matters, is really interpreting the theory at face value, and who is imposing reinterpretations that fit their agenda. So, if the idea is to make good sense of scientific discourse in general, to take it “at face value”, semantic realism might not be our best option after all.

As I said earlier, semantic realism can be held largely responsible for under-determination arguments, since it introduces a principled gap between truth and our epistemic abilities. This is what motivates invoking problematic modes of inference, such as inference to the best explanation, as analysed in Chap. 6. Arguably, the pessimistic meta-induction presented in Chap. 7 also arises because of semantic realism. A pragmatist notion of truth does not entail the same kind of ontological discontinuity between successive theories, since various ontologies can be considered equivalent if they have the same practical implications. In light of all this, we would be much better off if we abandoned the problematic (and unhelpful) thesis of semantic realism, at least as far as scientific theories are concerned.

³This kind of analysis could match the notion of categorical equivalence that has been recently proposed (Barrett and Halvorson 2016).

⁴Of course, this would deserve more analysis, in particular assuming that the models licensed by a theory evolve with time, which could be a difficulty for understanding theory equivalence. See Sect. 3.4.1 on theory identification.

8.4 A Pragmatist Alternative

According to Peirce (1931, 5.565), “Truth is that concordance of an abstract statement with the ideal limit towards which endless investigation would tend to bring scientific belief.” Another way to express this idea is to claim that a belief is true if it would withstand doubt, were we to inquire as far as we fruitfully could on the matter. In sum, pragmatic truth can be associated with a notion of ideal success. A complementary account associates truth with norms of assertion and inquiry: asserting something (and so, claiming that it is true) commits us to providing reasons in support for our claims and being held accountable for its implications, for instance (see Misak 2007).

Restricting our consideration to the context of science (which, according to Peirce, is continuous with everyday inquiry), there are reasons to think that the modal understanding of empirical adequacy proposed in this book is fit for being identified with theoretical truth. Empirical adequacy is a notion of ideal success, and it is associated with the aim of science, and thus plays a normative role for inquiry, just like pragmatic truth. It is also a modal notion, which fits well with the pragmatist emphasis on interventionist aspects, as explained in Sect. 4.2.4. The modal empiricist accepts that there are possibilities in the world, and this can also help make sense of the conditionals used in pragmatist accounts of truth (“it *would* withstand doubt”).⁵ If a theory correctly accounts for all possible manipulations and observations we could make in a domain of experience, then it would indeed “withstand doubt, were we to inquire as far as we fruitfully could on the matter”, and believing that a theory is empirically adequate is nothing more than believing that it would do so.

Another reason to identify empirical adequacy with truth stems from the strong analogies between the account of representation used in this book and aspects of philosophy of language.

There is a tight association between the notions of truth and meaning. Meaning is often understood in terms of truth conditions: the meaning of a statement is identified with the set of conceivable states of affairs that would make this statement true, also called its *intension*. This notion of meaning applies to particular utterances

⁵This modal aspect seems to be one of the reasons why van Fraassen does not endorse a pragmatist notion of truth, and opts for a distinction between belief and acceptance instead. He writes: “Whether belief that a theory is true, or that it is empirically adequate, can be equated with belief that acceptance of it would, under ideal research conditions, be vindicated in the long run, is another question. It seems to me an irrelevant question within philosophy of science, because an affirmative answer would not obliterate the distinction [between belief and acceptance] we have already established by the preceding remarks. (*The question may also assume that counterfactual statements are objectively true or false, which I would deny.*)” (van Fraassen 1980, p. 13, my emphasis). Contrarily to van Fraassen, I do not think that the question mentioned is irrelevant, because as argued earlier, his notion of acceptance appears to be very close to a pragmatic notion of belief, so accepting the latter does obliterate the distinction between acceptance and belief. And of course, I am not bothered by the idea that counterfactual statements are objectively true.

in context, but not necessarily to sentence types. The truth conditions of a sentence like “I am hungry” depend on who utters the sentence, so the sentence has no truth-condition outside of a context of use. However, its meaning can be captured by means of a character, which is a function from context to intension (Kaplan 1989): in each context, the sentence is true if the speaker is hungry.

The two-stage account of epistemic representation presented in Chap. 3 is a direct transposition of these accounts of meaning into a model-based conception of scientific theories. A model, interpreted in context, is the analog of an utterance, and its “meaning” or cognitive content is given by an interpretation, which determines its accuracy conditions (Sect. 3.2.2). This corresponds to the notion of intension. An abstract theoretical model presented outside of any context of use is analogous to a sentence type, and its “meaning” is given by a function from context to interpretation, which corresponds to a character (Sect. 3.3.3).

We could push the analogy between scientific and linguistic representation further, and observe that the laws and principles of a theory, what I have called meta-norms of representation, are quite analogous to grammatical rules, in that they constrain model construction in the same way grammatical rules constrain the construction of sentences. One could understand the theory of evolution as providing a “grammar” for representing a class of biological processes, and the formalism of quantum theory as a “grammar” for representing physical phenomena. The grammatical rules of natural languages are not purely conventional or arbitrary: they serve an aim, which is efficient communication, and not any rule can serve this aim (think of the way tenses allow us to talk about time). So, the fact that theoretical laws are revisable and responsive to experience rather than purely conventional is no objection to this comparison.⁶ As for theoretical terms, my suggestion is to identify their meaning with the functional roles that they play in representation, as captured by experimental and theoretical norms.⁷

In sum, the account of scientific representation provided in Chap. 3 is also an account of the “meaning” or cognitive content of scientific theories, and from a pragmatist perspective, there is no need to invoke any kind of realist interpretation

⁶Admittedly, the analogy between science and natural languages breaks down at some point, because the aim of scientific models is not efficient communication, but efficient experimentation. This might imply that theoretical laws are a bit more substantial than grammatical rules.

⁷Brandom (1994) has proposed to understand the meaning of linguistic terms in terms of introduction and elimination rules. The former correspond to the conditions in which it is appropriate to introduce the term, and the latter correspond to the inferences to which we commit by introducing the term. This account could possibly be adapted to our framework, with experimental norms playing the role of introduction rules and theoretical norms that of elimination rules. Note that this stance is compatible with a form of externalism about meaning, since norms are not necessarily perfectly known by the users of the theory. The correct application of a theoretical term can be deferred to the community. It could even be deferred to a *future* potential community, or to an ideal state of knowledge, assuming that these norms are not always clearly specified and can be improved upon. This is the case notably when these norms are tacit rather than explicit and rest on practical knowledge, as is often the case for experimental norms. This provides another response to Kripke and Putnam’s arguments mentioned in the previous section.

on top of this account. Assuming the traditional connection between truth and meaning, we can say, accordingly, that an interpreted model is “true” or veridical if it is accurate. At the level of theoretical models, we can understand truth in terms of ideal success: a theoretical model is “true” if its interpretations are accurate in all possible contexts. Finally, if theories are capable of being true or false, which is suggested by the way the term “theory” is commonly used, we can say that a theory is true if all its models are. These notions of truth correspond exactly to our definitions of empirical adequacy for models and theories (Sects. 4.2.6 and 4.3.3). I think that this notion of truth, combined with the idea that theories categorise contexts in certain ways, and are therefore disposed to evolve in particular ways when extended to new domains of experience, does justice to the way scientists identify equivalent theories. If two theories are applied in the same way in all contexts, support the same empirical inferences in these contexts, and are extended in the same way to new domains, then they are equivalent.

From this pragmatist perspective, modal empiricism becomes the position according to which the aim of science is to produce *true* theories, and science is generally successful in this aim. Our best theories are at least approximately true. So, modal empiricism is a form of realism if realism is understood minimally as a commitment to theoretical truth, irrespective of one’s conception of truth. Of course, all the arguments of this book can be carried over to this new version of modal empiricism. Theoretical truth can be justified by induction; the position accounts for scientific practice as well as for scientific success without resorting to inference to the best explanation, and it does not fall prey to the pessimistic meta-induction.

This notion of pragmatic truth does not have the same implications as the one that a semantic realist would adopt. However, it offers exactly the same linguistic resources. The modal structure of interpreted models can easily be described as a causal structure if one adopts a counterfactual theory of causation, for example Woodward (2003)’s interventionist theory of causation. One could go as far as to talk about unobservable objects acting as causes and effects when referring to specific nodes in this structure, even when the associated symbols are not directly interpreted empirically. These nodes are associated with theoretical terms which play specific functional roles in the models of the theory, and theoretical models are associated with types of targets which can also have a theoretical name. If scientific theories are modally adequate, a superficial analysis of theoretical content in terms of kinds and properties associated with dispositions, causal relations and laws of necessity seems available.

This kind of interpretation can help make sense of scientific discourse, where causal talk is ubiquitous, and where the existence of unobservable entities is not questioned. It is a crucial advantage for modal empiricism that it can make sense of this kind of discourse quite directly (see the arguments of Sect. 5.2.3). Without it, identifying empirical adequacy and truth would be problematic. However, we have seen reasons to doubt that scientific discourse is metaphysically loaded, so one should not put too much ontological weight on this way of talking.

For a modal empiricist, the “causal relations”, “objects”, “properties”, “dispositions” or “laws” that can be read off of the structure of scientific theories do not

describe the fabric of the world. These terms do not have the ontological meanings usually attributed to them in metaphysics. This is because the modal structure of a model interpreted in context is ultimately a structure of possible manipulations and observations. The accuracy of an interpreted model does not depend exclusively on the way the target of representation is independently of the user. It rather depends on the success of the user's interactions with this target. Particular purposes limit the range of possibilities considered for the target (Sect. 5.2.1). This relativity entails that the causal structure that an interpreted model describes is not a structure of "real relations" (Sect. 7.3), and by extension, the nodes of the causal structure that are only interpreted in terms of their positions within this structure are not "real objects". The types associated with theoretical models are more appropriately understood as types of contexts or activities than as natural kinds (see Sect. 3.3.3), and the functional roles played by theoretical terms only make sense relative to these types. But these differences in interpretation between modal empiricism and scientific realism do not appear at the surface level of scientific discourse. They only appear when one starts discussing semantic or metaphysical issues.

8.5 Scientific Objectivity

I wish to emphasise that the anti-realist stance presented here should not be associated with relativism, nor with a denial of scientific objectivity. The pragmatist notion of truth *is* objective. Scientific objectivity is sometimes understood in terms of context invariance, or in terms of value neutrality. Ideal norms of representation have both features: they should be reliable in all contexts, and therefore, they should not depend on local purposes and values. This is the case for norms of experimentation, which ensure the cross-contextual stability of experimental outcomes, as well as for theoretical norms.

Now, if scientific theories are objective, or tend towards objectivity, what prevents a realist interpretation?

There are two aspects that block an inference from objectivity, thus understood, to reality. The first one is that the kind of objectivity that is practically attainable is not necessarily absolute. We have seen that local purposes have the effect of limiting the range of possibilities considered in a given situation: some possibilities (including possibilities of discrimination) are excluded by stipulation because they are considered irrelevant in context (which can imply physical controls on the represented object). This could result in the fact that some patterns of relative necessity are "seen" in a context, but cannot be generalised to all contexts, which is a form of subjectivity. The fact that a norm is adequate in a large variety of contexts means fewer limitations for the range of possibilities considered, and more objectivity. However, this does not mean no limitations at all. As observed in Sect. 7.4, even if scientific theories embrace a large variety of potential contexts, there is no sense in which we can pretend having access to all possible contexts,

and our representations ultimately depend on our cognitive constitution and on our situation in the universe.

Now, even if theoretical models could embrace an unlimited range of possible contexts and thus achieve an ideal form of objectivity, I would still refrain from identifying the notion of objectivity implied with correspondence to mind-independent facts. The second reason to resist this move is that this particular notion of objectivity is actually orthogonal to the fact–value dichotomy. It rather follows the concrete–abstract dichotomy.

Concrete, contextual uses of theories are in general performative, and not only descriptive, in the sense that applying a model in context requires appropriate controls of the target, and that the inferences that the model affords guide the actions of its users when interacting with the target of representation. Contextual representational activities generally impact the world. For this reason, model choices in context can be guided by local values associated with particular purposes. In comparison, abstract representation, associated with the construction of what I have called general models in Chap. 3, is an “armchair” activity, and it does not have the same direct impact on the world. However, this does not mean that abstract representation is purely descriptive. We could rather say that if contextual representation is directly descriptive and performative, abstract representation is *indirectly* descriptive *and* performative, with the mediation of concrete representation.

An interpreted model is directly performative in the sense that it tells its users how they *should* interact with the target of representation. A general model or a theory is indirectly performative in the sense that it tells its users how they *should* represent target systems. It affects our representations rather than the external world. This impact is not local, but communal, so the process of abstraction, by which we gain objectivity, can indeed be associated with a form of value neutrality, or perhaps with a focus on abstract values that are more likely to be widely shared than local ones. The most abstract values might be identified with epistemic values. For this reason, I tend to agree that the ideal of value neutrality entertained by some philosophers in order to distinguish between theoretical and applied science makes sense to some extent, at least as a regulative ideal.⁸ However, abstracting away from local values is not the same as adopting a purely descriptive approach, and this does not warrant a realist perspective. Embracing more potential contexts actually brings us further away from local facts as much as it brings us further away from local values. Therefore, objective truth should not be equated with correspondence to the facts, but with ideal instrumental reliability when handling local purposes and local facts, whatever the purposes and facts involved. This blocks any tentative interpretation of scientific laws in terms of factive laws of nature, or in terms of dispositional essences of real entities.

⁸On this debate on scientific objectivity and the “value-free ideal”, see for instance Ruphy (2006), Elliott and McKaughan (2014), Longino (1996).

As we can see, one of the main characteristics of the pragmatist stance outlined here lies in its understanding of abstraction. An abstract representation, such as a scientific theory, does not float freely in a Platonic space. Its content can only be interpreted in terms of its potential applications. In this sense, indexicality is a defining characteristic of abstraction. This has implications for the interplay of values and facts: an abstract representation is no less “fact-free” than it is “value-free”, and performative and descriptive aspects are potentially mixed at any level of abstraction. The fact that abstract representation is grounded in concrete representation also means that the process of abstraction does not free us from being situated in this universe, nor from having a certain perspective on this universe associated with our cognitive constitution. Or to say it more pedantically, the ascent from the particular to the general that characterises abstraction is intensional rather than extensional.

As a consequence, adopting this pragmatist stance and claiming that scientific theories are objectively true if they are empirically adequate does not have any of the implications that a realist interpretation would afford. Nevertheless, it can directly account for causal discourse or discourse about unobservable objects, thanks notably to the commitment to modalities that characterises modal empiricism. It actually gives us exactly the same linguistic resources as a realist interpretation, at least if we remain at the level of scientific discourse and refrain from speculating about the fundamental nature of reality.

This leads us to questioning the status of metaphysics, a theme to which we now turn, by way of conclusion for this book.

8.6 Modal Empiricism and the Role of Metaphysics

The pragmatist stance endorsed by modal empiricism is an attempt to bring practical relevance to the way scientific theories are interpreted, by connecting this interpretation to practical uses. This stance could be adopted reflexively: what is the practical relevance of modal empiricism? In so far as modal empiricism purports to make sense of scientific practice and discourse, it cannot really have any impact on the way science is done, at least when this practice and discourse does not overlap with philosophy. I would say that endorsing modal empiricism should mainly affect the way we inquire into metaphysical questions.

For a modal empiricist, there is no reason to endorse a realist interpretation of a theory in so far as it does not make any difference in the way the theory is used. Scientific theories do not really inform us about what the world is like independently of us. Does it mean that metaphysical discourse is futile, or is there still room for metaphysics?

Metaphysics has not always been understood realistically. Kant viewed it as being about the conditions of understanding. Carnap viewed it as a matter of language choice. Others view it as a matter of conceptual clarification, which

is a rather neutral qualification.⁹ Assuming that a notion of pragmatic truth, conceptualised on the model of modal empirical adequacy, applies not only to scientific theories, but also to metaphysical discourse, the traditional metaphysical project which consists in inquiring about the fundamental nature of reality must be revised, but it does not necessarily mean abandoning all metaphysical discourse. It only means that metaphysics must be connected to experience in one way or another.

To be sure, interpreting metaphysical claims by means of pragmatic truth is a revisionary project, in the same way that Ladyman and Ross (2007, ch. 1)'s proposal of a "radically naturalistic metaphysics" is revisionary. Many metaphysical disputes are so disconnected from empirical considerations that it does not make much sense to claim that they will be resolved "at the end of inquiry". Ladyman and Ross do not think that all metaphysical claims are worthy of consideration. On the basis of Kitcher (1989)'s unificationist account of explanation, they propose that only the metaphysical claims that are capable of unifying two or more scientific hypotheses should be considered seriously. They also assume a principled "primacy of physical constraints", so that at least one of the hypotheses they unify should come from physics. Besides, they explicitly exclude "projects that are primarily motivated by anthropocentric (for example, purely engineering driven) ambitions" (p. 36). The notion of unification they employ matches quite well with the approach I advocate, given the important role that the notion of unification plays for modal empirical adequacy, hence for pragmatic truth. However, this account appears to be quite restrictive from a pragmatist perspective. It seems to ignore that there is a rich life outside of theoretical science, without which science would probably be meaningless. Putting contextual practice at centre stage implies that this rich life should not be construed as derivative and devoid of interest for metaphysical inquiry.

Let me suggest a way of doing justice to one of the motivations of Ladyman and Ross's naturalistic metaphysics, which is to connect metaphysics to empirical considerations, in a less restrictive way. What I suggest is to apply a pragmatist methodology, and to examine the function played by various metaphysical concepts in philosophical inquiry, so as to propose a pragmatic reinterpretation of these concepts, in the same way causal discourse was interpreted in the previous section. This reinterpretation should be based on the previously mentioned idea that abstract concepts are indexical: they must be interpreted in terms of their potential application to concrete representational uses. However, because of their metaphysical nature, they will not be tied to particular domains of experience or to particular types of activities. They will enjoy a greater level of generality. This is the unificatory aspect. I suspect that this pragmatic reinterpretation of metaphysical concepts will generally reduce these concepts to reflexive principles that regulate our way of representing the world in general, as will be illustrated shortly. Science has a special place in this project, because it aims at pragmatic truth. However, this place is not necessarily exclusive.

⁹Guay and Pradeu (2020) propose a taxonomy of various approaches.

This project is undoubtedly revisionary. A pragmatic reinterpretation of metaphysical debates is unlikely to satisfy most metaphysicians engaged in traditional projects. However, I think that it has the capacity to endow these debates with more tangible implications.

Let me illustrate this approach with the debate on free will and determinism. Entertaining an idealistic notion of free will that is opposed to the idea of metaphysical determination is very problematic from a pragmatist perspective, since it is not clear that the concept has practical implications. On what grounds shall we say that a particular decision was free in the metaphysical sense? However, the concept of free will can be elucidated in terms of its functional relations to other, more tangible concepts, such as moral responsibility, associated with social norms, or the absence of external constraints on one's decisions. This is precisely the approach followed by compatibilists. According to some compatibilists, we have free-will in so far as we are able to do what we wish without impediments in our way. They argue, on the basis of this interpretation of the concept, that free will is actually compatible with determinism: whether our wishes are or are not determined by an underlying reality is simply irrelevant.

This interpretation of free will is perfectly compatible with a pragmatist stance. The compatibilist notion of free will has potential applications in concrete representational uses: it can easily be interpreted in context. However, one could go even further, since the concept of determinism is also problematic for a pragmatist. We do not have epistemic access to putative laws of nature. So, the concept also needs to be reinterpreted in a tangible way, and a natural reinterpretation is one in terms of ideal predictability: a system is deterministic, in a pragmatist sense, if an ideal empirical knowledge of the system by an external agent at a given time allows for perfectly accurate predictions of any relevant property of the system in its future.

Assuming this reinterpretation, we can positively assert, on the basis of quantum mechanics for instance, that physical systems are not always deterministic. "Hidden variables" that would restore a deterministic interpretation of quantum mechanics at the metaphysical level are by definition excluded, because they are not accessible to external agents. However, *some* systems are deterministic. Remember that in our account of scientific representation, physical systems are identified relative to a context, assuming variables of interest and standards of precision, and some isolated systems, such as the ones represented in Newtonian mechanics, can be considered deterministic in the pragmatist sense, because given relevant degrees of precision, the properties of the system can be predicted with perfect accuracy. Coarse-graining a situation, that is, lowering standards of precision, actually ensures that one will reach a deterministic system at some point. So, the question "Is free will compatible with determinism?" is not settled by a compatibilist reinterpretation of free will, since this notion of free will might well be compatible with metaphysical determinism, but it could still be incompatible with our pragmatist reinterpretation of it. The question becomes: "Is it possible, for an external agent, to predict systematically and with perfect accuracy the decisions of a person when this person does what she wishes without external impediments?" Standards of accuracy should of course be adapted to the kind of decisions that interest us.

This question has more practical relevance than the original one. It seems to matter whether our decisions can or cannot be predicted in principle by external agents, since this predicability can be associated with a valuable notion of privacy. It also has interesting ramifications, because being able to predict the decisions of a person with perfect accuracy seems to give us a greater ability to impede the actions of this person. The answer to this question could even have social implications. And as far as I know, this is an open question. If living organisms are chaotic systems, the answer could be negative, because measuring the initial state of an organism with sufficient precision for prediction would require controls that would impede the freedom or integrity of the organism, for instance, or because it would require isolating the organism in a way that impedes free decisions. It is *a priori* conceivable that the functioning of complex organisms, and their agency in particular, rests on a certain “privacy” of internal states that is disrupted by external measurements, as a matter of (situated) necessity. If this was the case, it would make sense to define agency in terms of this notion of privacy. The answer to our question lies in life sciences and cognitive sciences, but also in an epistemological analysis of the notion of predictability, which is a more reflexive aspect (however, no hypothesis from physics seems required).

Note that if the decisions of a person can in principle be predicted with perfect accuracy by an external agent, then the concept of free will understood in terms of the idea that a person “could have done otherwise” is undermined, since this kind of predictability implies that there is a model of the person where only one possible decision for the person is represented. If this model is empirically adequate, then there is no possible situation where the person does otherwise. So, this reformulation of the debate is not entirely foreign to traditional philosophical analyses of free will, and it could explain why some have the impression that the compatibilist solution dodges the real issue instead of addressing it. We can see that the modal component of modal empiricism plays a valuable role here, because it allows us not to fall back on an impoverished interpretation of metaphysical concepts.

The main advantage of this approach is that we have traded a somehow intractable metaphysical problem that could only be answered *a priori*, if at all, for an empirical question. We have empirical means of identifying persons who act according to their wishes, and we have theoretical and empirical resources to inquire into a potential relation of (situated) necessity between free human agency and unpredictability, or at least we might have them in the future. Furthermore, as just said, our commitment to natural modalities allows for a relative proximity with the traditional formulation of the debate, at least in some respects.

Let me present a second example: the debate on reduction and emergence. This debate concerns the possible causal autonomy of higher-level entities with respect to lower-level entities. Kim (2000) has proposed an influential, *a priori* argument, to the effect that assuming certain premises, such as supervenience, the causal closure of the lower level and the absence of overdetermination, emergence is impossible: the lower level does all the “causal work”, so to speak. However, this debate also needs to be reformulated, because according to a modal empiricist, causal attributions are always relative to a context of representation. A tangible question is whether the

model of a situation interpreted in a fine-grained context always supervenes on a model of the same situation using a coarse-grained context by means of systematic rules connecting the two. The idea is, again, to connect the metaphysical notion of reduction to potential representational uses.

Assuming this reformulation of the debate, Kim's argument is not necessarily applicable. Having a low-level description of a target system, for example, a physical description, implies having a context where low-level (physical) properties are empirically accessible. From a pragmatist perspective, representation is not passive. Making these properties accessible generally requires experimental interventions. In the case where we could show that observing higher-level properties, for example, biological properties, is incompatible with observing their lower-level constitution (because it would disrupt the system), it would make sense to talk about emergent properties, simply because a fine-grained and a coarse-grained context are mutually exclusive. There could be cases where making accessible the exact position of microscopic particles composing a system could not be done without disrupting the higher-level causal structure of the system, which would mean changing the context of representation. In such a case, the high-level description of this causal structure does not supervene on any low-level description, or only on a coarse-grained one, thus blocking Kim's argument. A metaphysician could entertain the idea that there is an underlying causal structure that explains the higher-level one, but the higher level does not reduce to the lower level in the pragmatist sense, and we could talk about practical emergence. Whether there are such cases is an empirical question (quantum mechanics could have this kind of implication for entangled systems). Again, we have turned an a priori issue into one that is answerable by empirical means.

I leave it as an exercise to the reader to apply the same approach to other metaphysical debates, for example in the philosophy of time (where I strongly suspect that Putnam (1967)'s relativist argument against presentism becomes inapplicable) or concerning the interpretation of quantum mechanics (where the pragmatist stance could be similar to Healey (2012)'s approach).

This is a crude presentation, since my purpose here is not to dwell on these metaphysical questions, nor to engage in empirical speculations. The main message of these illustrations is that a careful pragmatist analysis makes it possible to connect metaphysics with experience, and that the results of metaphysical inquiry, once focused on practical aspects, might be significantly different from what we obtain by adopting a traditional metaphysical stance. This approach blurs the separation between metaphysics and science, in the spirit of naturalised metaphysics, because metaphysical questions become partly empirical. However, as the illustrations show, even when they are reformulated, metaphysical questions are not *purely* empirical. They require that we reflect on our representations and on the relations between various contexts of representation. They demand that we clarify the relationship between our concepts and experience, or that we establish links between various concepts, including concepts, such as predictability, which are "meta-empirical". This reflexive aspect could be the main characteristic of metaphysics, and since the

connection with experience plays a crucial role in pragmatic reformulations, one could say that a pragmatist approach also blurs the distinction between metaphysics and epistemology.

I doubt that a traditional metaphysician will be totally satisfied with these reformulations: she would say that we are now answering different questions. It is also likely that some metaphysical questions are not susceptible of being reinterpreted in this way, because they are embedded in an approach that is at odds with a pragmatic stance. Philosophical thought experiments that call for us to imagine other metaphysically possible worlds do not make much sense in this approach. Can we conceive a world populated with “phenomenal zombies”, that is, people who are like us in all physical respects but have no phenomenal consciousness (Chalmers 1996)? In our account of representation, someone else’s phenomenal consciousness cannot be represented pragmatically, since it is, by definition, inaccessible to external agents, so the representation of a zombie world and the representation of our world would be identical. On the other hand, applying this representation could imply that the user of the representation has phenomenal consciousness while being located in that zombie world, which contradicts the premises of the thought experiment. In any case, this kind of thought experiments cannot tell us anything about what is possible or not *in this world*, since the notion of possible world has no practical implication. So, no conclusion follows.

Perhaps modal empiricism lacks resources to address “deep” metaphysical questions. But we could be suspicious about the relevance and soundness of these questions and associated thought experiments.¹⁰ The description of a zombie world *seems* to make sense on the surface of it, but are we really capable of representing such a world? Does it not rest on a dubious semantics?

My verdict in this matter is that we should be careful with abstract representation. Abstraction should not be construed purely descriptively and a-contextually, but also normatively and indexically: an abstract representation is a norm for possible concrete representations. Situated representation, potential or actual, should always take centre stage in philosophical analysis. Good concepts tell us how we *should* think in particular contexts. I strongly suspect that some metaphysical questions, which have all the appearances of legitimate conceptual analyses, only occur because we take our abstract representations too seriously, and forget their normative dimension with regard to concrete representation. The purpose of a pragmatist stance is to remind us that this normative dimension is always present. Because inapplicable norms are useless, this dimension implies that some ways of asking metaphysical questions are more *valuable* than others: the ones that have practical implications. This defiance of unrestrained metaphysics is a general characteristic of empiricist positions (and it is also shared, to some extent, by Ladyman and Ross’s naturalistic metaphysics), but I believe that modal empiricism and the pragmatist stance associated with it, when

¹⁰This is a point where modal empiricism in philosophy of science agrees with the homonymous position in the epistemology of metaphysics. See Sect. 1.3.

compared with traditional non-modal versions of empiricism, put epistemic agents in a less passive position that is better apt to serve as a useful guide for philosophical inquiry.

The main obstacle to adopting a pragmatist stance could be that it apparently clashes with a naturalistic world-view. This, of course, depends on how naturalism is understood. The pragmatist stance does not oppose the idea that philosophy is continuous with science, which is a traditional theme of naturalism. Quite the contrary actually, as we have seen. However, it requires abandoning the idea that there is a “view from nowhere”, that is, that there can be representational content without any tie to an epistemic agent, not even an abstract (indexical) one (this perspectivist stance that opposes the “view from nowhere” is shared by other authors who call themselves realists, such as Massimi (2018) and Giere (2010)). There is no reason to think that epistemic agents are “unnatural”, and they can be represented by other epistemic agents, but it is true that the representation relation itself is not entirely naturalised in this approach. This could be seen as a problem. I see it as a mark of modesty (after all, attempts to naturalise intentionality or values have not been particularly fruitful), and I am convinced that it is actually an advantage once one starts running into metaphysical troubles. Thinking in this way forces us to be relevant. And we *ought* to be relevant. This could be the most important message of modal empiricism.

References

- Barrett, T., & Halvorson, H. (2016). Morita equivalence. *Review of Symbolic Logic*, 9(3), 556–582. <https://doi.org/10.1017/s1755020316000186>.
- Belot, G. (2012). Quantum states for primitive ontologists. *European Journal for Philosophy of Science*, 2(1), 67–83.
- Brandom, R. B. (1994). *Making it explicit: Reasoning, representing, and discursive commitment* (vol. 183). Cambridge: Harvard University Press.
- Buckareff, A. (2010). Acceptance does not entail belief. *International Journal of Philosophical Studies*, 18(2), 255–261. <https://doi.org/10.1080/09672551003677838>.
- Chalmers, D. (1996). *The conscious mind: In search of a fundamental theory* (vol. 4). Oxford: Oxford University Press.
- Coffey, K. (2014). Theoretical equivalence as interpretative equivalence. *The British Journal for the Philosophy of Science*, 65(4), 821–844. <https://doi.org/10.1093/bjps/axt034>.
- Dummett, M. (1978). *Truth and other enigmas* (vol. 31). Cambridge: Harvard University Press.
- Elliott, K. C., & McKaughan, D. J. (2014). Nonepistemic values and the multiple goals of science. *Philosophy of Science*, 81(1), 1–21. <https://doi.org/10.1086/674345>
- Giere, R. (2010). *Scientific perspectivism*, paperback edn. Chicago: University of Chicago Press.
- Guay, A., & Pradeu, T. (2020). Right out of the box: How to situate metaphysics of science in relation to other metaphysical approaches. *Synthese*, 197(5), 1847–1866. <https://doi.org/10.1007/s11229-017-1576-8>.
- Healey, R. (2012). Quantum theory: A pragmatist approach. *British Journal for the Philosophy of Science*, 63(4), 729–771.
- Horwich, P. (1991). On the nature and norms of theoretical commitment. *Philosophy of Science*, 58(1), 1–14. <https://doi.org/10.1086/289596>.

- Israel-Jost, V. (2015). L'observation scientifique: aspects philosophiques et pratiques. No. 8 in *Histoire et philosophie des sciences, Classiques Garnier*, Paris.
- Kaplan, D. (1989). Demonstratives: An essay on the semantics, logic, metaphysics and epistemology of demonstratives and other indexicals. In J. Almog, J. Perry, & H. Wettstein (Eds.) *Themes from kaplan* (pp. 481–563). Oxford: Oxford University Press.
- Khrentzos, D. (2016). Challenges to metaphysical realism. In E.N. Zalta (Ed.) *The Stanford encyclopedia of philosophy*, winter 2016 edn. Stanford: Metaphysics Research Lab, Stanford University.
- Kim, J. (2000). *Mind in a physical world: An essay on the mind-body problem and mental causation*. London: MIT Press
- Kitcher, P. (1989). Explanatory unification and the causal structure of the world. In P. Kitcher, W. Salmon (Eds.) *Scientific explanation* (vol. 8, pp 410–505). Minneapolis: University of Minnesota Press.
- Kripke, S. (1980). *Naming and necessity*. Cambridge: Harvard University Press.
- Ladyman, J., & Ross, D. (2007). *Every thing must go: metaphysics naturalized*. Oxford: Oxford University Press.
- LaPorte, J. (2004). *Natural kinds and conceptual change*. Cambridge: Cambridge University Press.
- Longino, H. (1996). Cognitive and non-cognitive values in science: Rethinking the dichotomy. In L.Hankinson Nelson, & J. Nelson (Eds.) *Feminism, science, and the philosophy of science* (pp. 39–58). Kluwer Academic Publishers.
- Maher, P. (1990). Acceptance without belief. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, 1990*, 381–392.
- Massimi, M. (2018). Perspectival modeling. *Philosophy of Science*, 85(3), 335–359. <https://doi.org/10.1086/697745>.
- Maudlin, T. (2018). Ontological clarity via canonical presentation: Electromagnetism and the Aharonov–Bohm effect. *Entropy*, 20(6), 465. <https://doi.org/10.3390/e20060465>.
- Misak, C. (2007). Pragmatism and deflationism. In C. Misak (Ed.) *New pragmatists* (pp. 68–90). Oxford: Oxford University Press
- Mitchell, S. (1988). Constructive empiricism and anti-realism. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, 1988*, 174–180.
- Nguyen, J. (2017). Scientific representation and theoretical equivalence. *Philosophy of Science*, 84(5), 982–995. <https://doi.org/10.1086/694003>.
- Peirce, C. S. (1931). *Collected papers of Charles Sanders Peirce*. Cambridge: Harvard University Press.
- Putnam, H. (1967). Time and physical geometry. *Journal of Philosophy*, 64(8), 240–247. <https://doi.org/10.2307/2024493>.
- Putnam, H. (1975). The meaning of 'Meaning'. *Minnesota Studies in the Philosophy of Science*, 7, 131–193.
- Reich, E. S. (2011). Finding puts brakes on faster-than-light neutrinos. *Nature*. <https://doi.org/10.1038/news.2011.605/news.2011.605>
- Ruphy, S. (2006). Empiricism all the way down: A defense of the value-neutrality of science in response to Helen Longino's contextual empiricism. *Perspectives on Science*, 14(2), 189–214. <https://doi.org/10.1162/posc.2006.14.2.189>.
- Ruyant, Q. (2020). Semantic realism in the semantic conception of theories. *Synthese*. <https://doi.org/10.1007/s11229-020-02557-8>.
- van Fraassen, B. (1980). *The scientific image*. Oxford: Oxford University Press.
- van Fraassen, B. (1989). *Laws and symmetry* (vol. 102). Oxford: Oxford University Press.
- Woodward, J. (2003). Making things happen: A theory of causal explanation. *Oxford studies in philosophy of science*. New York: Oxford University Press.