Synthese Library 402
Studies in Epistemology, Logic, Methodology, and Philosophy of Science

John Wright

An Epistemic Foundation for Scientific Realism

Defending Realism Without Inference to the Best Explanation

Springer

Synthese Library

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

Volume 402

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

John Wright

An Epistemic Foundation for Scientific Realism

Defending Realism Without Inference to the Best Explanation

Springer

John Wright
Department of Philosophy
University of Newcastle
Callaghan, NSW, Australia

Synthese Library
ISBN 978-3-030-02217-4 ISBN 978-3-030-02218-1 (eBook)
https://doi.org/10.1007/978-3-030-02218-1

Library of Congress Control Number: 2018957998

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

Contents

Chapter 1
Introduction: Realism and Reason

Scientific Realism with respect to a theory is, at least as a first approximation, the doctrine that the entities – including the unobservable entities – postulated by the theory exist and behave (more or less) as the theory says they do. It is an ontological or, perhaps, metaphysical thesis. A philosopher who is a Scientific Realist about, for example, a theory of electrons might hold that electrons exist and behave more or less as the theory says they do.

But, if we are to be *justified* in claiming that certain unobservable entities exist, and that they behave more or less as our theory says, then it seems we must have *good reasons* for these claims. And so, a problem arises: Given that electrons, or quarks, or mesons, are unobservable, how can we have good reason to believe they exist? Further: if it is the case that any scientific theory will be underdetermined by the actual data on which it is based, how can we have good reason to believe these entities conform to *our* accepted theory rather than some alternative that provides a different explanation of the same phenomena? More generally, how is possible for us to have *good reasons* for the claims that Scientific Realists make? This book is an attempt to answer those questions.

Although Scientific Realism is itself an ontological or metaphysical doctrine, there are epistemological questions that, at least on the face of it, would need to be addressed if it is to be claimed we have *good reason* to believe Scientific Realism. Our concern is with those epistemological questions. The first chapters of the book defend realism against arguments purporting to show we cannot have such good reasons, the later chapters develop and defend an account of the nature of the reasons for realism.

In the Chaps. 2 and 3 we consider, and rebut, general arguments that *prima facie* seem to show that we cannot have good reason for Scientific Realism. In Chap. 4 we critically examine what is perhaps the most influential argument *for* Scientific Realism: that it provides us with the best explanation of the "the success of science". It is argued that there is a lacuna in this argument that has as yet not been satisfactorily resolved. The lacuna is that we do not – as yet at any rate – possess good reason for saying that if a theory is the best it is therefore true or likely to be

J. Wright, *An Epistemic Foundation for Scientific Realism*, Synthese Library 402, https://doi.org/10.1007/978-3-030-02218-1_1

true. An alternative route to Scientific Realism, which does not rely on the principle of Inference to the Best Explanation (IBE), is therefore needed.

Such an alternative route is outlined in Chap. 5. More specifically, in that chapter an account is developed of that which entitles us to say unobservable entities *exist*. In Chap. 6 an answer is sketched to the question: "What entitles us to say that these unobservable entities conform to one set of laws rather than another?". This chapter relies on earlier work by the author. In Chaps. 7 and 8 the approach is applied to some examples from the history of science.

1.1 An Outline of the Argument of the Book

We noted in the previous section that there are general arguments that seem to challenge the idea we have good reason to believe in Scientific Realism. Perhaps the most fundamental of these is Humean skepticism about induction.

Humean skepticism about induction can be interpreted narrowly or broadly. Interpreted narrowly, it purports to show that induction – construed as, for example, enumerative induction – cannot be rationally justified. Construed broadly, it purports to show ampliative inference generally cannot be rationally justified. It will here be assumed that "Hume's Argument" against induction is to be construed broadly: if valid, it undermines not just enumerative induction but ampliative inference generally. And, of course, any inference from observations to a Scientific-Realist claim *will* be ampliative. It will be ampliative in at least two ways: it will go beyond that which we have observed in telling us that *unobservable* entities exist, and it will also go beyond that which we have observed in telling us those entities obey certain universal laws. So, Humean skepticism about induction, if justified, would appear to undermine the claim we can have good reason for Scientific Realist claims.

Hume's Argument against induction does not seem to figure very prominently in contemporary discussion of the epistemic challenges confronting Scientific Realism. Part of the reason for this, no doubt, is because Hume's Argument is not *exclusively* a problem for Scientific Realists: it is a problem for anyone who wishes to make claims that go beyond that which we have observed. But I suspect there is another reason. Many philosophers, I suspect, are inclined to simply put Hume in the "too hard" basket. The prospects for finding a solution are so slim that many, perhaps, feel their time would be better spent working on other things. However, there is reason to think this defeatist view might be unwarranted. In recent years there has, I think, been progress in moving towards a solution to the problem of induction.[1] These themes are explored in Chap. 2.

---

[1] More specifically, the work of David Stove and Laurence BonJour have given us reason to be optimistic about the problem of induction. The work of Stove and BonJour is discussed in the next chapter.

Chapter 2 has two main aims. The first is to argue that general skepticism about the rational justifiability of ampliative inference is unwarranted. More specifically, it is argued that there some statements that are clearly synthetic but for which we can have good, if defeasible, *a priori* reason. It is argued that the existence of such statements shows we may be able to develop a justification of ampliative inference generally. The second aim of the chapter is the more specific one of developing and defending a justification of *enumerative* induction. Quite apart from constituting a reply to more general forms of skepticism, the conclusions defended in Chap. 2 will also aid us in more specific ways. One of the main claims of this book is that a type of inference that leads to *realism* turns out to be very closely related to enumerative induction. More specifically, it will be argued, enumerative inductive inferences and certain types of inferences to unobservable entities are but different forms of an underlying form of inference. Developing a rational justification of enumerative induction is a key step in the development of our account of the nature of the reasons for realism.

Chapter 3 is concerned with arguments purporting to show we cannot have good reason for Scientific Realism.

Perhaps the most prominent in recent discussion of these sceptical arguments comes from the "Pessimistic meta-induction on the History of Science". The history of science reveals to us that past theories have reasonably frequently turned out to be false: Does this not make it rational for us to suppose that our theories, too, will eventually be discovered to be false, or at least undermine our confidence that they are true? The force of the Pessimistic Meta-Induction depends, very roughly, on just "how much" of our falsified theories have been shown to be false. It is argued that key examples from the history of science indicate that, although the Pessimistic Meta Induction does not undermine all realist claims, it does cast in to doubt some *particular* realist theses. Specifically, it casts in to doubt the idea that if our sole reason for believing in the *existence* of something is that it is postulated by our best theory, we thereby have good reason to believe it exits. Directed against existence claimed, the Pessimistic Meta-Induction does seem to have some force. It is argued in later chapters that the particular perspective developed here does enable us to deal with this challenge to realism.

There are of course other challenges to realism. A range of these is considered later in Chap. 3. One challenge comes from the fact that any given, finite body of data can be explained in many, different ways. We will here refer to this as the underdetermination of theory by *actual data*.[2] If this form of underdetermination is

[2] Perhaps the expression "underdetermination of theory by data" is most usually used to refer to the Quinean thesis that there will be more than one theory capable of explaining "all possible data". Nonetheless, it seems to the present author to be useful to speak of the underdetermination of theory by actually obtained data. This is the thesis that, given any actually obtained body of data D, there will be a number of possible ways of explaining the data. This is to be distinguished from the thesis that, given any actually obtained body of data D, there will be a number of incompatible ways of using induction to make predictions about what we will observe in the future. Conceivably, there could be two different theories T and T* that have the same observational consequences, but disagree about what is going on at the theoretical level. If so, it seems natural to describe this as a

granted, the question arises: what reason do we have for accepting one particular theory rather than any one of the others capable of explaining the same data?

The whole of this book can be regarded as a response to the issues raised by the thesis of the underdetermination of theory by actual data. Some general responses to the problem of underdetermination are given in Chap. 3, but a more detailed response is developed in Chap. 6. In that chapter the notion of the independence of theory from data is introduced. The degree of independence of a theory is, roughly, a measure of its lack of *ad hoc*-ness. It is argued that if one theory is more independent of the data than its rivals then we have more reason to believe that its *empirical predictions* will turn out to be correct. It should however be noted that the fact that a theory is highly independent of the data does not, by itself, give us reason to believe that the theory is true *tout court.* In particular, it does not give us reason to believe the parts of the theory about unobservable entities are true. So, the notion of the independence of theory from data does not furnish us with a *complete* reply to the challenge to Scientific Realism from the underdetermination of theory by data. (The reply is "completed" by the notion of independence working in conjunction with other notions, to be described below).

Some more problems for the rationality of belief in Scientific Realism considered in Chap. 3 are: The problem of "equivalent descriptions", an argument from Bayes' Theorem purporting to show that the probability of a theory must always be zero, the argument from the "Experimentalists Regress", arguments appealing to the theory laden-ness of observation and the argument from unconceived possibilities.

In Chap. 4 we discuss the influential idea that realism is justified because it provides the best explanation of the success of science. But, as we have noted, it is controversial whether the fact that a theory is the *best* shows it is likely to be *true*.

As noted, Scientific Realism, at least as a first approximation, consists of two claims:

(i) The entities, including the unobservable entities, postulated by scientific theories *exist*.
(ii) The entities, including unobservable entities, postulated by scientific theories *behave* more or less as the theories say they do.

We might refer to these two aspects or dimensions of Scientific Realism as the *existence* aspect and the *behaviour* aspect. If we are to have good reason to believe Scientific Realism with respect to some theory T, we must, it seems both have good reason to believe that the entities it postulates exist and that they behave as T says they behave. The main aim of Chap. 5 is to argue that, at under certain circumstances, we can have good reason to believe that the unobservable entities postulated by a theory *exist*. The key concept used is that of an "Eddington Inference". Given the centrality of this concept to the overall position of this book, it is appropriate to give a brief outline of it here.

---

case of underdetermination by actual data. Also, such a case need not be an example of underdetermination by all possible data: once all possible data are in it might turn out that both T and T* have been falsified.

The term “Eddington Inference” is an allusion to the scientist Arthur Eddington. In his *The Philosophy of Physical Science* Eddington asks us to imagine an ichthyologist investigating the size of fish. All the fish he studies have been caught in a net with holes two inches across. The size of two inches has been chosen “blindly”, that is, in ignorance of the size of fish in the sea. The ichthyologist inspects the fish in the net and notes that there are no fish in it of less than two inches. Ought the ichthyologist to conclude that there are (probably) fish in the sea of less than two inches?

Eddington himself – rather surprisingly – argues that the ichthyologist ought *not* to draw this conclusion. He defends this view on grounds that we might see as possibly influenced by neo-Kantianism, or perhaps the last sentence of Wittgenstein’s *Tractatus*, or even, perhaps, as an early statement of something resembling Hilary Putnam’s “internal realism”.[3] Roughly, and briefly, Eddington says anything not capable of being caught by the net of our observational or conceptual apparatus cannot constitute a part of science, and so we ought to remain silent about what it may or may not be like.

In this book we will not be concerned with Eddington’s neo-Kantian or “internal realist” perspective – if that is what it is. Rather, here we will accept what I would assume to be the response of common-sense to the ichthyologist’s catch: if the catch contains fish of a range of sizes from two inches upwards, and holes in the net are exactly two inches, then we have good reason to think that there probably *are* fish in the sea shorter than two inches. And one reason for this is quite straightforward: if there were no fish in the sea smaller than two inches, the “blindly chosen” size of the holes in our net would happen to have coincided with the size of the smallest fish in the sea. Since this is unlikely, we have reason to believe there are probably fish in the sea smaller than two inches. The point can, of course, be generalised. Are there entities too small to see with the unaided senses, or with our most powerful observational apparatus? If not, then a highly improbable fluke would have occurred: the size of the smallest entities *there are* would happen to have coincided with the size of the smallest entities detectable by our apparatus. Since this is surely unlikely, we have reason to believe smaller entities probably exist. The inferences to smaller fish, or to entities too small to see with our apparatus, are all examples of what will be termed “Eddington-inferences”.

One of the main aims of Chap. 5 is to argue that Eddington-inferences do not have the same problem as inferences to the best explanation. More specifically, it is argued that Eddington-inferences can be given a justification similar to the justification of induction defended in Chap. 2. Eddington inferences do *not* require us to assume, for example, that simpler theories are more likely to be true. But, it is

[3] For a discussion of Eddington’s Kantianism, see Arthur Ritchie *Reflections on the Philosophy of Arthur Eddington*, Cambridge University Press, 1948, pp. 2–5. Ritchie sees Eddington as influenced by Kant, but also by a variety of idealists, including Berkeley. For a somewhat different suggestion concerning the similarity between Eddington and Wittgenstein’s *Tractatus*, see John D. Barrow “The Mysterious Lore of Large Numbers” in *Modern Cosmology in Retrospect*, by B. Bertolotti et al. (eds) (Cambridge University Press, 1990), pp. 67–93, esp. p. 73 and note 7.

argued, an Eddington-inference to an existence claim does increase the probability that the existence claim is true.

We noted above that there seem to be two dimensions to Scientific Realism: the existence dimension and the behaviour dimension. Even if we have good grounds to believe that entities of a certain sort exist, it does not follow we have good grounds for saying they behave in a particular way. Perhaps one theory T can explain the data pertaining to some class entities; but, given the underdetermination of theory by actual data, so may many other theories. What reason do we have for saying that the entities to which we are led by an Eddington-inference obey the laws of T rather than the laws of some other theory? On the view developed in Chap. 6, it is argued that the notion of *the independence of theory from data* can provide such a reason. If the entities to which we are led by an Eddington-inference may, as far as the data indicate, obey either T or T*, but T is more independent of the data than T, then we have probabilistic reason to believe it is more likely that the entities obey T. And so: we have reason for a Scientific Realist view of theory T.

The reader may recall that we earlier stated that the fact that some theory T provides the best explanation of some data does not give us good reason to think that the unobservable entities postulated by the theory exist. And yet, it has just been claimed that the fact that a theory T is more "independent of the data" does give us reason to prefer T to other theories. There is, however, no contradiction here. No claim is made here that the notion of the independence of theory from data, *by itself*, warrants Scientific Realism with respect to a theory. But a key claim of this book is that an inference to the most highly independent theory, *working together* in a certain way with an "Eddington-inference", *can* justify some scientific realist claims.

Appeal to Inference to the Best Explanation (it will be argued) is confronted with a problem: we do not seem to be in possession of a *justification* of the claim that if a theory is the best it is therefore more likely to be true. But the approach advocated here is, it is claimed, free of that difficulty. The inferences used here, it is argued, *can* be given (probabilistic) justifications. The thesis will be defended that theories that are more independent of their data have a higher probability of empirical success (but *not* a higher probability of truth) than less independent theories. It will also be argued we have probabilistic reason to accept the conclusions of Eddington inferences. So, on the view advocated here, some scientific realist claims *can* be given probabilistic justifications.

In the final chapters of the book, the account developed is applied to examples from actual science. In Chap. 7, the arguments of Maxwell, Einstein and Perrin in support of the existence of atoms are examined. In Chap. 8, we examine inferences to the existence and broad features of those parts of the universe lying beyond that which we can observe. In these final chapters it is argued that both claims about things too small to be observed, and claims about things too distant to be observed, can be give probabilistic support of the kind developed here.

Chapter 2
The Skeptical Arguments Against Realism I: Inductive Skepticism

Our aim is to show how we can have good reasons for claims about the unobservable, theoretical claims made by science. But: there are objections to the thesis that we can have such reasons. The aim of the early chapters of this book is to critically evaluate those objections. In this chapter we will only be concerned with one of the objections: Humean skepticism about induction.

2.1 Why a Reply to Humean Skepticism About Induction Is Needed

Possibly, historically the single greatest cause for scepticism about the idea that science gives us good reason for its claims comes from Hume's Argument against induction.[1] Given the extreme and fundamental nature of the argument, it is perhaps surprising it does not occupy a position of more prominence in contemporary philosophical discussion. But, given the aims and specific approach of this book, we cannot here avoid it. Here we are concerned with the questions: "Do we have good reason to adopt a realist stance towards scientific theories?" and "If so, what are those reasons?" To adopt a realist view of a theory T is, roughly, to say the entities postulated by T exist and behave more or less as T says they do. So, to have good

[1] Hume gave his argument against induction in his *An Enquiry Concerning Human Understanding*, especially Section IV: "Sceptical Doubts Concerning the Operation of the Understanding". However, what philosophers nowadays generally have in mind when they refer to "Hume's Argument" perhaps goes rather beyond Hume's brief treatment of the topic. Some more recent statements of what has come to be known "Hume's Argument" are Brian Skyrms *Choice and Chance*, Chapter Two "The Traditional Problem of Induction" and Wesley Salmon *An Encounter with David Hume*.

In his *Popper and After*: *Four Modern Irrationalists* David Stove identifies Hume's scepticism about induction as the underlying source of an irrationalism about science he discerns in Popper, Kuhn, Lakatos and Feyerabend.

J. Wright, *An Epistemic Foundation for Scientific Realism*, Synthese Library 402, https://doi.org/10.1007/978-3-030-02218-1_2

reason to adopt a realist view of T we need to have good reason to believe the entities, including the unobservable entities, postulated by T *exist*, and that those entities *behave in a particular way*. On the approach advocated here, the existence of unobservable entities is established via what we here refer to as an Eddington inference. In Chap. 5 it is argued that Eddington inferences are to be justified in the same way that inductive inferences are to be justified. It is therefore important to show that the form of justification used is, in fact, *good*.

The other thing that needs to be done to justify realism about T is to show we have good reason to believe the entities postulated by T behave in a particular way. But to say that the entities postulated by T behave in a particular way is to say that they conform to a particular pattern or law(s). Once again, induction, or something like it, would be needed to show this.[2]

Finally, in the last chapter of this book, it is argued that the Cosmological Principle is needed if we are to have good reason to believe things about regions of space beyond the observable universe. But, it will also argued, the grounds for believing in the Cosmological Principle are also at least very similar to those for believing in induction. A justification of induction, and inferences very similar to it, lies at the core of the position developed here.

One thesis to be argued for, especially in Chap. 4, is that we do *not* possess a satisfactory justification of inference to the best explanation (IBE). So, here no attempt is made to justify realist claims with IBE. But we *do* attempt to justify realist claims with inductive inferences, Eddington inferences, the notion of the independence of theory from data and the Cosmological Principle. Clearly, if the approach advocated here is to be *preferred* to IBE-based approaches, then we require a satisfactory justification of the forms of inference used here. The aim of this chapter is to lay the groundwork for such a justification.

Amongst more recent philosophers who have attempted to solve the problem of induction are: David Stove,[3] F. J. Clendinnen[4] and Laurence BonJour.[5] One aim of this chapter is to defend and extend some ideas of these authors.

This chapter is divided in to three parts. The first part is an exposition and analysis of Hume's argument against induction. The second is a critical examination of a crucial assumption of (this interpretation of) Hume's Argument. It is argued that this crucial assumption is in fact false, and hence that Hume's Argument against induction fails. In the third part steps are taken towards developing a positive argument for induction. The ideas developed in this section have points of similarity, but also points of difference, particularly from those of Stove and BonJour.

---

[2] The notion that plays this role on the view given here is the notion of the independence of theory from data. But, the argument given in favour of that notion is, again, very closely related to the argument for induction.

[3] See for example David Stove *The Rationality of Induction*, 1986, Oxford, Clarendon Press.

[4] See, for example, F. J. Clendinnen "Rational Expectation and Simplicity" in *What*? *Where*? *When*? *Why*?; *Australasian Studies in History and Philosophy of* Science, vol 1, (1982), pp.1–25.

[5] See Laurence BonJour *In Defence of Pure Reason* (Cambridge University Press).

2.2 Hume's Argument

Hume's Argument purports to show that a justification of induction is not possible.[6] I will here assume that a "justification of induction" would need to show that we have good reason to believe induction will "work"; where, we can say, induction "works" if it at least takes us from true premises to true conclusions more frequently than would sheer random chance.[7]

It might perhaps be felt that such an account of what it is to justify induction really does set the bar very low: if induction takes us from true premises to true conclusions *only slightly* more often than does random chance, then, it may be protested, induction has surely only been "justified" in an at best very minimal sense. Two points can be made in reply. First, it will later be argued that we do not have even a minimal justification of this sort for any use of IBE sufficient to issue in a claim about the existence of unobservables. Second, it will be argued that there is good reason to believe the justification that can be given for Eddington inferences – which does issue in claims about unobservables – is in fact *rather stronger* than the corresponding justification for induction. These two points together give us reason to prefer Eddington inferences over IBE as a route to realism about unobservables. So, although the standards for a justification of induction might be seem to be rather low, it will be argued that they are – given our aims – all that is required from a justification of induction.

Let us now briefly review Hume's Argument. A justification of induction of induction would, presumably, need to be some sort of an *argument*. But what type of argument might it be? Plainly, an inductive argument will not do since it is induction we are trying to justify: using an inductive argument would beg the question.

Might we justify induction using a deductively valid argument? It is a feature of deductively valid arguments that the conclusion can assert no more than is explicitly or implicitly asserted by the premises. The conclusion of a "justification of induction" – that induction will lead us from true premises to a true conclusion more often than would sheer chance – is evidently not an analytically true proposition. It is contentful, contingent and synthetic. Moreover, it will evidently need to be in some sense "about the future", since we surely want to have good reason to believe induction will work *in the future*. So, if the conclusion is to say no more than is implicitly contained in the premises, then the *premises* of any deductively valid justification of induction will need to be synthetic, contingent and, in some sense, "about the future". But then the question arises: What reason would we have for accepting such premises? If they are "about the future" we evidently cannot establish them merely by observation. It would seem some sort of *inference*, perhaps from what we have observed, would be required to give us good reason for the premises. If that inference were an inductive inference, we would be back to begging the question. And if

---

[6] Hume's own statement of what has since come to be known as "Hume's Argument" can be found in Hume's Enquiry Concerning Human Understanding, paragraph 4.

[7] This account of what is required of a justification of induction follows Skyrms.

the premises were to be established by some ampliative method (method “J”, let us say) other than induction then we would seem to need some reason for trusting method J. Presumably, method J would stand in need of justification as much as does induction itself. So: the question would arise: How is J to be justified? Since J is also ampliative, the questions of justification that would arise for J would be essentially the same as those for induction. And so, we find ourselves back at square one.

The only other option – at least for a deductively valid justification of induction – would seem to be to try to justify induction from *a priori* premises. But this approach is also confronted with difficulties. The only things we can know *a priori,* it seems very plausible to say, are analytic truths that do not make contentful claims about the world. If so, *a priori* claims would certainly not make contentful claims about future states of the world. But the conclusion of a justification of induction plainly would need to make a contentful claim about future states of the world: it would need to tell us that induction will work in the future. Since the conclusion of a deductively valid inference cannot tell us anything not already implicitly contained in the premises, the conclusion of such an inference, with *a priori* premises, cannot make any contentful claim about the future. Consequently, such an inference could not constitute a justification of induction.

It might perhaps be suggested that induction might be justified by an argument that is neither inductive nor deductively valid: call it an inference of type K. But, as we noted above, the question would then arise: What reason would we have to trust inferences of type K? If we were to attempt to justify inferences of type K by induction, we would, again, be begging the question. The difficulties that confronted us in attempting to justify induction with deduction would also confront any attempt to justify K with deduction. And, of course, if we were to attempt to justify K using K itself then we would quite clearly again be guilty of begging the question. Perhaps we could justify K using some type of inference L. But then the question would arise: What reason would we have for trusting L? And so the problem goes on…

The above is, I take it, a more or less standard exposition of what has come to be known as Hume’s Argument against induction.

2.3 Analysis of Hume’s Argument

In this section a key assumption of Hume’s Argument will be identified and critically evaluated. One assumption of the argument is that *a priori* propositions can only be analytic. This assumption is clearly necessary for Hume’s Argument, as presented above, to go through. But in this section, it will be argued it is false.

In considering the claim that an *a priori* proposition cannot be synthetic it is useful to distinguish between *a priori* knowledge and *a priori* reasonable belief. Knowledge is usually taken to be belief that has at least the properties of being true and being justified. It is also customary to add some extra property or properties to deal with the Gettier examples.[8] But, for a proposition to be *a priori reasonable*, it is clearly not required that it also count as *knowledge*. There is, obviously, no requirement that it in fact true. Neither is there any requirement that it be able to deal with Gettier examples. And, plausibly, to require that there be *some a priori* reason to accept a proposition is rather weaker than the requirement it be "justified". So: showing there can be *a priori* synthetic reasonable belief would appear to be a much easier task than showing there can be *a priori* synthetic knowledge.

But now, in order to have a justification of induction – at least on the account offered here – it is not clear that the premises of such a justification must count as *a priori knowledge*. At least on the view offered here, we have a justification of induction if we have *good reason to believe* that induction will work. And, in order to get that conclusion, it would not seem to be necessary to start from premises that constitute *knowledge*. It *may* be possible to arrive at our desired conclusion if we merely have good reason for our premises.

Still, it may be protested, whether what we require are examples of synthetic *a priori* knowledge or synthetic *a priori* reasonable belief, we still require the impossible. We cannot, it may be protested, have any reason at all to believe a synthetic claim unless we have some empirical evidence for it. But, if we do have some empirical evidence for it, the claim is evidently not *a priori.*

The idea that we cannot have any reason to believe a synthetic claim unless we have some *empirical* evidence for it does have strong intuitive appeal, and we can perhaps bring out its intuitive appeal in the following way. Imagine a man sealed in a box. Inside the box he receives no messages from the outside world. He can, of course, speculate about the nature of the world outside the box. But, without any kind of causal input or perception coming from the outside world, his thoughts will be *nothing more* than speculations. He might, for example, speculate that there might be a house, or a tree, or a flower outside the box, or nothing at all outside the box, but as long as he is causally isolated from the outside, there can be no epistemic basis to these speculations.

The idea expressed in the above paragraph expresses, I think, a common Empiricist Intuition. In this section it will, however, be argued there is good reason to think it is wrong. But, if we are to reject this Empiricist Intuition, it seems that we also ought to be able to explain *why* it seems so compelling. An explanation is offered later in this section.

The following arguments against the Empiricist Intuition resemble, but also slightly differ from, arguments developed by Laurence BonJour.[9]

One difficulty for what we are here calling the "Empiricist Intuition" is that, given any body of empirical data, there can be a number of different, and incompatible, ways of accounting for that data, and yet it can seem intuitively clear that we do have *more reason* to believe some of these accounts than others. But if there is more reason to believe, say, T_1 than T_2 even though there is just as much *empirical* reason for one as there is for another, it would seem to follow that, in some sense, at

[8] Edmund Gettier "Is Justified True Belief Knowledge?". Analysis, 23(6), pp.121–123, (1963).

[9] BonJour, *op cit*, pp.1–6.

least some component or aspect of the reason for T_1 must be non-empirical. And if that is so it would appear to follow there must be such a thing as a non-empirical reason.

We can illustrate this by modifying the above example. Suppose a window were opened on the box allowing the man to see out. We would surely be inclined to say the man now *could* have reason to believe something about what was outside his box. If he were to see a tree through the window, he might thereby have reason to believe there is a tree out there. But: that there is a complication here is only too familiar to philosophers. What would give the man reason to believe it was a *tree* out there, rather than something else that merely caused the visual sensation of a tree? What gives him reason to believe there is anything at all other than the mere visual sensation of the tree? That the man may in fact have good reason to believe there is a tree out there, causing his experiences, need not be disputed. But, that there is something "out there" *causing* his visual experiences is, it would seem, *not* given by those visual experiences all by themselves. Empirical input is, perhaps, necessary for the man to have good reason to believe there is a tree out there, but it is at least philosophically controversial that it is by itself sufficient. And so, a challenge arises for the empiricist: If we grant that the man *could* have good reason to believe the tree exists, how is it possible, within empiricism, to explain how the man has good reason for *this* belief rather than one of its empirically or experientially equivalent alternatives?

There is of course a tradition, coming from the later Wittgenstein, that asserts we do not have what is properly called "knowledge" of our own private, mental states: the expression: "I know (or have good reason to believe) that is an X" can only be correctly used when the X is a publically observable object of some sort. But even if this is granted, direct observation, all by itself, seems to fall far short of giving us all that common-sense would say we *do* have good reason to believe. At any one time, all the *empirical* evidence that we have will be compatible with a number of incompatible and in some cases highly implausible hypotheses about *the future*. For example, all the empirical evidence we currently have is compatible with the hypothesis that the material objects we are now perceiving will in the year 2100, or on "D-Day", all cease to exist, or that they will all inexplicably turn blue or start to jump and down. A similar bizarre hypothesis, not about the material objects themselves but about our perceptions of them, might also be advanced. We surely want to say we have *more reason* to believe that, in the future, things will proceed more or less as they do now rather than that on some specified "D-day" they will all suddenly turn blue, be jumping up and down and so on. But if all the *empirical* evidence so far is compatible with both common-sense and the bizarre hypotheses, it would appear our reason *now* for preferring the common-sense hypothesis must have at least a non-empirical component. Examples can also be given of hypotheses that are flagrantly irrational but empirically indistinguishable from what we believe to be the truth.[10] We surely *do* have reason to believe certain hypotheses over their

[10] Hypotheses attributing non-standard metrical properties to space are a source of examples of this sort. Possibly the example of this kind most strongly in conflict with what we believe to be true is

experientially or (up until now) empirically equivalent rivals, but if this is so, it surely follows that at least a *part* of the reason for some hypotheses must be non-empirical.

It might perhaps be protested that it simply does not follow from the fact that two hypotheses have (up until now) been empirically indistinguishable that we do not now have *any* empirical reason to prefer one to the other. One hypothesis might be preferable to the other because it is, for example, simpler, where we also have *general* empirical grounds for preferring (for example) simple hypotheses.

However, it is not too difficult to see this suggestion runs in to a difficulty. What sort of *empirical* evidence might we have for the preferability of (for example) simple theories? A natural suggestion might be along the following lines: "A survey of the history of science shows that (for example) simple theories have been better at leading to successful prediction (or some other form of confirmation) than theories that were not simple, and so we have good inductive grounds to believe in this particular case the simpler theory is better." But in the present context it is far from clear this really would give us good grounds for preferring simple theories. One initial point: the justification plainly relies on induction, and we have yet to see how that is to be justified empirically. But there is another difficulty. If there are a number of theories that can all account for the data so far, then there will also be a number of theories that also have a good track record. For example, all the theories asserting that something bizarre will start to happen on "D-day" will, so far, have had a track record every bit as good as those theories we regard as being much more rationally credible. These theories will not be as *simple* as the others, they will have some other property which we will call "*schimplicity*". The assertion that schimple theories are good will have received just as much empirical confirmation as the assertion simple theories are to be preferred. And so, it remains unclear how purely empirical considerations could supply us with reason to prefer the hypothesis we regard as correct with rivals asserting something bizarre will happen on D-day. But

---

the concave Earth hypothesis developed by Mostapha Abdelkader: "A Geocosmos: Mapping Outer Space into a Hollow Earth" in *Speculations in Science and Technology.* (vol 6, pp.81–89, (1983)). For a discussion, see "Quine on space-time" by J. J. C. Smart, and the reply by Quine, in *The Philosophy of W. V. Quine* edited by Lewis Hahn and Paul Schilpp (Open Court, 1986). Abdelkader develops a view on which the surface of the Earth is concave, and in which the sun, stars and galaxies are contained within the sphere that is the surface of the Earth. It appears that by modifying the paths taken by rays of light, other motions and the laws of physics, Abdelkader's view can be made empirically indistinguishable from the view we believe to be true. Surely, Abdelkader's view is *flagrantly* unreasonable. However, if it is unreasonable even though empirically equivalent to our view, whatever reason there is to favour our view must be non-empirical.

In his discussion of Abdelkader's hypothesis, Martin Gardner allows that empirical data could not give us reason to prefer our view to Abdelkader's. Gardner instead argues that Abdelkader's view must be rejected because of its great complexity. While this is may be (at least in the present author's view) correct, it leaves unanswered the question: What, if anything, entitles us to assume simplicity is a sign of truth? In the present context the point to note is that it is far from clear how there might be *empirical* evidence that simplicity is a sign of truth. The system of physical laws Abdelkader postulates, in all their baroque complexity, will have every bit as good a track record as our laws. So, the empirical evidence could equally well be taken to support the hypothesis that simplicity is a sign of truth, *or* that baroque complexity is a sign of truth.

we surely do have more reason for the former, and some of it must be non-empirical reason.

Of course, a hard-line empiricist might simply embrace the consequence that we have no more reason for one theory than any of its empirically equivalent rivals. But, whatever might be said in defence of such a position, *it certainly does not seem to accord with our intuitions, or with common sense.* Our intuitions, surely, tell us we do have *more reason to believe* realism rather than idealism or that 5 min from now the furniture in this room will be more or less as it is now rather than blue or jumping up and down. So, appealing to our intuition here seems to lead us to say we can and sometimes do have non-empirical for preferring one belief to another.

In summary, our intuitions can lead us in opposite directions. On the one hand, when we consider the man in the sealed box, it does seem extremely implausible indeed to think he could have any kind of reason at all for any belief about how things are outside the box. But, when we consider some pairs of empirically equivalent but incompatible synthetic theories, our intuitions, I think, lead us in the opposite direction: they do, I think, lead us to the conclusion we do have some kind of non-empirical reason for preferring one synthetic proposition to another. A philosopher who rejected the possibility of non-empirical reasons on the grounds they are counter-intuitive would, it seems, be attending to only some of our intuitions while ignoring the consequences of others. The *prima facie* appeal of the Empiricist Intuition ought, it seems, *only* be seen as *prima facie*, and ought not to be taken as a decisive difficulty for the idea of *a priori* reasons.

The difficulty for the Empiricist Intuition can perhaps be put a slightly different way. The strict empiricist can, perhaps, be seen as presenting us with a problem: "*How could* we possibly know about anything "out there", without observing it or being causally affected by it in some way?" But the force of this question, as an objection to *a priorism*, starts to look a lot less once we realise that observation, by itself, doesn't get us very far in giving us knowledge either. *A priori* knowledge perhaps seems "mysterious", but so is any kind of knowledge that takes us beyond what we have actually observed. If we reject the former because it is mysterious, so it seems ought we to reject the latter. But to reject the idea that we know, or at least have good reason to believe, things like: the Sun will rise tomorrow, or my (currently unobserved) keys are in the drawer, and so on, is to reject a multitude of Moorean facts. We surely ought to accept that we can and do have reasonable belief that goes beyond that which we have experienced.

So, we now find ourselves now confronted with the question: *How* is it possible for a person to have a reason for, or a justification of, a belief that goes beyond that which they have observed? We consider one influential answer to this question in the next section.

2.4 Reliabilism

Of course, it is one thing to say a speaker can have reason for beliefs that go beyond what has been observed, but quite another to embrace some form of *a priorism*. One influential account of the nature of justified belief that does not require *a priorism* is reliabilism. I will take reliabilism to at least assert:

One sufficient condition for a belief to be justified is for it to be produced by a reliable method, where a method is reliable if and only if beliefs produced by it are always, or at least mostly, true.

An often-made objection to reliabilism is that it yields results in conflict with our intuitions about justification and reasonable belief.[11] The following imaginary case can be used to bring out the counter-intuitive nature of reliabilism. Let us imagine two reputable scientists, who we will call Sally and Wally. They are both working on devices that, they hope, will be able to read people's memories. The devices on which they are working scan a certain part of a subject's brain and then develop hypotheses about the content of the memories stored in that part of their brain. The devices then display on a screen some message such as "Subject A has a memory of injuring their ankle while skiing" or "Subject B has a memory of meeting the Queen of England", and so on.

Now let us suppose, although neither Sally nor Wally have yet got a device that actually works, their research is showing signs of promise: they both feel they might be on the verge of developing a device that can actually read memories. But, at this point, the way they further develop or refine their devices could go in two ways. We will refer to these possible refinements as Refinement A and Refinement B. Suppose Sally chooses to modify her machine according to Refinement A, and Wally chooses to modify his according to Refinement B. As far as Sally and Wally know, both refinements stand a reasonably good chance of working, but they do not have compelling evidence that either refinement definitely will work. Both of them, let us say, agree the two modifications have an equally good chance of working. Prior to the modifications being carried out, there is no reason, available to either Sally or Wally, to think one modification is more likely to work than the other.

Let now suppose that Refinement A adopted by Sally does in fact work. If Sally's (now refined) device tells her, for example, that "Subject X has a memory of meeting the Queen", then there really is a memory stored in X's brain of meeting the Queen. Wally's device, refined in way B, however, does not reliably work. Wally's

[11] See for example Laurence BonJour "Externalist Theories of Empirical Knowledge" *Midwest Studies in Philosophy* 5, 53–57. (1980). In this paper BonJour the case of "Clairvoyant Norman". BonJour argues that, despite the fact that clairvoyant Norman's beliefs are formed in a way that reliably leads to truth, we would not say he had good reason for his beliefs. The imaginary example of "Sally" and "Wally" in the text to follow is of the same general sort as BonJour's "Clairvoyant Norman", but it will also be argued this imaginary case has an advantage over the example of "Clairvoyant Norman" .

device might say that a subject has a memory of injuring their ankle even though the subject has no such memory.

Recall we are assuming that prior to the modifications, neither Sally nor Wally have reason to suppose one modification more likely to work than the other. Both Sally and Wally are in the same epistemic position. *After* the modification, the only difference between Sally and Wally is that Sally's device does in fact reliably produces truths, but Wally's does not. Neither of them at this stage have any independent confirmation of the fact that Sally's modification worked while Wally's did not. But would we say that if Sally *comes to believe*, for example, that her subject has a memory of meeting the Queen that Sally's belief is justified while the beliefs Wally comes to about his subject's memories are not justified? I don't think we would be inclined to say this. If, prior to the modifications being carried out, neither Sally nor Wally have any reason to suppose one modification is more likely to work than another, then I do not think we are inclined to say Sally's beliefs are *more justified* than Wally's.

Of course, subsequent investigations might provide independent confirmation that Sally's device is reliable while Wally's is not. We might for example, get independent confirmation that subject A once met the Queen, whereas the "memories" supposedly revealed by Wally's machine never receive any independent corroboration. But, prior to any such independent corroboration, I do not think we are inclined to say Sally's beliefs about her subject's memories are any more justified than Wally's belief about the memories of his subjects, even if they are reliably produced.

This would seem to present a problem for reliabilism. Sally's beliefs are produced by a reliable method, Wally's are not. Yet, it seems intuitively compelling that Sally's beliefs are no more *justified* than Wally's. Certainly, I think we are strongly inclined to say Sally has *no more reason* for her beliefs than Wally.

It is perhaps worth noting that there is a particular way of defending reliabilism that would not work here. It has been suggested that a belief produced by a reliable method need not be justified if that belief was produced by a method that has a "defeater".[12] Suppose, for example, a belief is produced by clairvoyance, in a situation in which we have no independent corroboration of the reliability of clairvoyance as a way of arriving at beliefs. But, the fact is, claims of clairvoyance are generally regarded with scepticism. If someone were to tell us, for example, that exercising their faculty of clairvoyance had told them a particular horse was going to win, we would be disinclined to believe their prediction was justified. That a particular belief is produced by clairvoyance is a *defeater* of the claim of the belief to be justified.

However, there does not seem to be any defeater present in the situation just described. Recall we are assuming that Sally and Wally are both reputable scientists and both feel the device on which they are working might be on the verge of working successfully. So, in this context, the assertion "The machine is successfully

[12] See Alvin Goldman *Epistemology and Cognition*, Harvard University Press, (1986), pp.111–112.

detecting the memories of subjects" would not seem to be an implausible or hard-to-believe assertion. The fact that Sally's beliefs are produced by the machine would not seem to defeat that the claim that Sally's beliefs are justified.

Let us now modify our example a little. Suppose Sally and Wally now use their machines on the same subject. Sally's machine says the subject has a memory of meeting the Queen, Wally's machine says the subject has no such memory. Reliabilism would seem to tell us that what Sally's machine says is justified, while what Wally's machine says is not. But would Sally and Wally, under these circumstances, be able to tell which one was justified? Of course, they may be able to by other ways, such as by asking the subject, or investigating their history. But in this case, the sheer fact that the assertion produced by Sally's machine is produced by a reliable method is not enough for either Sally or Wally to *tell* that it is justified.

These considerations seem to bring out a way in which reliabilism defeats the whole purpose of having a notion of "good reasons". What is "the point" of a notion of rationality, or of good reasons? It is plausible, I think, that one of the things we want a notion of rationality to do is to guide or *help* us to the truth. We cannot see directly how various less accessible parts of the world are. We cannot see directly how things were a long time ago, or in the future, or at very remote points in space; neither can we directly see how things are at the level of the very small. One of the things we want rationality to do is to guide us to the truth about these less accessible regions. We cannot see directly in to, say, the interior of the atom, but one of the things we hope rationality will be able to do is to guide us to the truth about what is going on inside the atom.

A parallel case might be helpful here. Suppose we are crossing a desert and come to a point where we can go either left or right. One way will lead us to water, the other way will not, but we do not know which is which. However, at this point we also encounter two *guides*. One of the guides is reliable, the other is not. If we pick the right guide, we will be able to get to the water. If we pick the wrong guide, we will not. So, if we can *tell* which of the guides is reliable, we are better off having encountered the guides. But if it is just as hard to pick the right guide as it is to pick the right path through the desert, then we are no better off having encountered the guides. The sheer fact a reliable guide exists is of no help to us. We need to be able to identify the reliable guide. For the existence of the guides to be useful to us, it must at least be the case that it is *easier* to tell which guide is reliable than it is to tell which path will get us to the water.

If rationality is to be able to play the role of guiding us to the truth, then, pretty clearly, we must be able to *tell* whether or not it is rational to believe a particular proposition. Good reasons for a belief, or a justification of that belief, must themselves be *accessible* features of a belief. That we have good reasons for P would surely need to be a more accessible fact than the fact that P is true. But this feature of justification seems to be lost if we accept reliabilism. That the beliefs to which we are led by Sally's machine are justified would seem to be *as inaccessible as* the fact that they are true. And this is would seem to defeat the purpose of having a notion of justification, or good reasons.[13]

In the opinion of the present author, examples such as this show that reliabilism leads to implausible and unattractive results.[14] It is natural to suppose that perhaps one reason why people sometimes nonetheless stick with reliabilism, despite its *prima facie* counter-intuitive nature, is because it is felt that the consequences of abandoning it would be even worse. More specifically, it is perhaps feared that abandoning it might commit us either to scepticism or to embracing the idea that we can have *a priori* reason for synthetic propositions, and both of those options might seem even less welcome than any counter-intuitive features reliabilism may have. However, it will here be argued that this is not so. In *some cases*, it will be argued, the synthetic *a priori* can be quite acceptable.

2.5 Synthetic a Priori Reasonable Belief

There are, to be sure, many synthetic propositions of which it would be absurd to suggest we might have *a priori* knowledge. For example, no one, presumably, would seriously suggest we might know *a priori* that Mount Everest was about 29,000 feet high. To assert that people could know this *a priori* would certainly seem to attribute to them some mysterious faculty. But, it will be argued, there are *some* synthetic

[13] The argument just given, although based in part on Bonjour's "Clairvoyant Norman", is designed to bring out how reliabilism has difficulty in accommodating the idea that a function of rationality is to act as guide to truth about the less accessible parts of reality. The argument given also uses ideas developed by Keith Lehrer and Cohen "Justification, Truth and Coherence", *Synthese*, 55, pp. 191–207. Lehrer and Cohen consider Descartes' "evil genius" hypothesis. They argue that Descartes' hypothesis creates a difficulty for reliabilism. More specifically, they argue that a person misled by Descartes' evil genius would have just as much reason as us to believe in the existence of an external, material world. But, of course, the method by which the person deceived by Descartes evil genius arrives at their beliefs does not reliably produce truths. Hence, the way in which a person deceived by Descartes' evil genius constitutes a counter example to reliabilism.

It is suggested that perhaps the argument given here might have an advantage over that given by Lehrer and Cohen. It is open to dispute whether the person deceived by Descartes' demon, or a brain in a vat, would have any beliefs about material objects at all. This has been argued, for example, by Hilary Putnam. If this is correct, the brain in a vat does not falsely believe the *proposition* "There are material tables.", rather, when the brain in a vat says to itself the *sentence* "There are material tables", the sentence does not refer to material tables. If this is granted, it becomes unclear whether the brain in a vat really does have a vast number of false beliefs. Presumably the same may hold of a person deceived by Descartes' genius. Perhaps, if the brain's beliefs are interpreted as being merely about its own experiences, its beliefs might be true. And so it seems to be at least an arguable thesis that the brain in a vat is not a counter-example to reliabilism at all. But the argument presented here would not appear to be confronted with that difficulty. The brain in a vat argument does not cast in to doubt the either the meaningfulness or the falsity of Wally's beliefs.

[14] For a critical discussion of BonJour's position, see for example Jose Zalabardo "BonJour, Externalism and the Regress Problem", *Synthese*, 148, 1, pp.135–169. (2006)

propositions which are such that, to say we can have *a priori* reason to believe them, does not seem at all to involve attributing to us any kind of "mysterious faculty".[15] They are, it will be argued, when properly understood quite innocuous.

It is appropriate to briefly outline the relation the position defended in this section has to some of the main positions on the status of the *a priori*. Following BonJour,[16] we may say there seem to be at least the following four main positions:

(i) Traditional rationalism: if the mind knows *a priori* P then the mind grasps not only that P is true but P is necessarily true; moreover the knowledge that P is true is certain and empirically indefeasible.
(ii) Moderate empiricism: The only truths knowable *a priori* are analytic truths and tautologies.
(iii) Radical empiricism: There can be no *a priori* reasons for any proposition.
(iv) Moderate rationalism: It is possible for intellectual insight to furnish us with *a priori* reasons, not just for analytic truths but also for synthetic truths; however, these *a priori* reasons may be empirically defeasible.

The position defended here is a form of Moderate Rationalism, although it is (in a way to be explained) also very close to Moderate Empiricism.

It would take us too far from our central concerns to engage in a full discussion of all four of these options. But it seems to the present author that the criticisms of Traditional Rationalism and Radical Empiricism given by BonJour are effective.[17]

---

[15] This is not to deny that we still need an understanding of how we can come to know, for example, analytic truths *a priori*. But such cases seem far less "mysterious" than would *a priori* knowledge of something like "Mount Everest is over 29,000 feet high."

[16] See BonJour, *In Defense of Pure Reason*, (Cambridge University Press, 1998), pp.15–19.

[17] BonJour criticizes traditional rationalism by arguing that there are cases of propositions that certainly seemed to be supported by *a priori* insight, but which we now believe to be not just defeasible, but empirically defeasible. One example is Euclidean geometry. We now believe, not only that the axioms of Euclidean geometry *might* be false, but that there is empirical evidence that in the vicinity of heavy bodies such as the Sun they are *in fact* false.

BonJour also rejects the radical empiricism of Quine. For Quine, the evidence for apparently *a priori* assertions, such as those of logic and mathematics, is ultimately purely empirical. The assertions of logic and mathematics, for Quine, lie very close to the centre of our "web of belief". As such, they are not vulnerable to falsification by any empirical observations we might make: only assertions close to the observational periphery are vulnerable to such falsification. Michael Devitt suggests that this apparent immunity to empirical falsification is the reason why we are (mistakenly, in Devitt's view) inclined to regard the statements of logic and mathematics as *a priori*.

For BonJour, we do in fact have some positive reason to believe the laws of logic and mathematics. How might a Quinean do justice to this? For a (Quinean) radical empiricist, any evidence that accrues to the laws of logic and mathematics lying in the centre of the web must ultimately come from the observational periphery. But Quine also says the interior of the web is underdetermined by its observational periphery. The actual content of the interior of the web is determined, not just by empirical input, but also by considerations of simplicity and conservatism. But now it seems: if we are to have good reason to believe the contents of the interior of the web, we must also have some sort of reason to think the simplicity, for example, of our web is an epistemic point in its favour. What kind of reason might we have for this? It seems it cannot be purely empirical reason, for the reasons given in the previous section of this chapter. So BonJour concludes a radical empiricist Quinean cannot do justice to the idea we have reason for the (empirically undeterdeter-

If Traditional Rationalism and Radical Empiricism are rejected, we are left with Moderate Empiricism and Moderate Rationalism. There is at least one point on which Moderate Empiricists, Moderate Rationalists (and even Traditional Rationalists) agree:

It is possible to have *a priori* reason for analytic truths.______(AA)

Both the Moderate Empiricists and the Moderate Rationalists will want to add to (AA). The Moderate Empiricists will add that we cannot have *a priori* reason for anything other than analytic truths and tautologies, while the Moderate Rationalists will say we can. Even Traditional Rationalists will accept (AA), provided that the *a priori* reason is said to be indefeasible. Only the Radical Empiricist will reject (AA).

There is of course the question of *how* we can have *a priori* reason for analytic truths. The present author has nothing on say on that question. But the fact that (AA) is widely, although not universally accepted makes it for certain purposes desirable. Our overall aim is to justify induction. To do this, it has been argued, we need some *a priori* reason. If we can produce such reasons using nothing more than (AA), then our justification of induction may be acceptable to a wide variety of philosophical positions.

But, of course, it is natural to immediately protest this appeal to (AA) is not of much help if our aim is to justify induction. (AA) only licenses us in asserting analytic truths *a priori*. But, as we have noted, analytic truths aren't sufficient to justify induction. To do that, we need *synthetic a priori* truths. So – even if (AA) does represent common ground between Moderate Empiricists, Moderate Rationalists and Traditional Rationalists – it certainly *seems* clear it couldn't be sufficient to give us a justification of induction.

The main thesis to be defended in the next section is that actually it *is* possible to arrive at some synthetic *a priori* reasonable beliefs using no more than (AA) as a starting point.

Of course such a feat might seem impossible: no one can get the rabbit of synthetic truth out of a purely analytic silk hat: if it seems as though this has been done then, it might be thought, we must have been deceived by some kind of trickery. So, it is worthwhile here giving some brief preparatory remarks about how we will proceed. At the core of the argument to be presented, there is a distinction between two types of ampliative inference. These two types of ampliative inference are *contentfully ampliative* inferences and *epistemologically ampliative* inferences. If an inference is contentfully ampliative, the conclusion has *content* that goes beyond the content of the premises, but when a speaker knows the conclusion is true, they *need not know anything more* than when they know the premises are true. However, if an inference is *epistemologically* ampliative, a speaker who knows the conclusion to be true thereby *does know more* than a speaker who merely knows the premises to be true.

---

mined) contents of the interior of our web. So, the radical empiricist Quinean cannot do justice to the idea we have reason for the laws of logic and mathematics.

Epistemologically ampliative inferences stand in need of justification just as much as do inductive inferences. However, contentfully ampliative inferences that are *not* epistemologically ampliative are different. Such inferences do not stand in need of justification in the way that epistemologically ampliative inferences do. It will be argued that these inferences ought to be acceptable to the Moderate Empiricist, and that they require commitment to no more than (AA). Finally, and crucially, it will be argued these inferences are sufficient to furnish us with reasonable synthetic *a priori* beliefs, and with a justification of induction.

2.6 Examples of Synthetic a Priori Reasonable Beliefs

In this section some examples of synthetic but *a priori* reasonable beliefs will be given. First we need to introduce the notion of a "blindly chosen" observation. Suppose a coin has been tossed a very large number of times. Sometimes it will have come up Heads, other times, Tails. We select, from amongst this very large number of tosses, some tosses to observe and record. This selection is made *blindly* if, prior to choosing a particular tossing to observe and record, we have no reason to believe one outcome (Heads or Tails) is any more likely than the other outcome. Then we may surely assert:

If, hypothetically, a coin had come up heads every time in one hundred blindly chosen tosses, then there would be good reason to believe the coin is not fair; more specifically, there would be good reason to believe that the propensity for Heads to come up is greater than the propensity for Tails to come up._____(1)

Note that (1) does not tell us that, if the coin has been tossed, it *does have* a propensity to come up Heads – it only says this is, to some degree, a reasonable thing to believe. Surely (1) has at least some *a priori* plausibility. But, it is, perhaps, not very plausible to claim that (1) is synthetic. Perhaps (1) is merely an analytically or conceptually true claim about what it is, under certain circumstances, rational or reasonable to believe. Certainly, no claim will be made here than (1) is synthetic. Of course, if a philosopher were to claim that that (1) *is* synthetic, while also allowing it has some *a priori* plausibility, they would be granting that it is possible to have *a priori* reasons for synthetic propositions and we would have already established our desired conclusion. So, here it will be *conceded* that (1) is analytic.

But, although (1) is, quite plausibly, *not* synthetic (and will here be assumed to be not synthetic) the following claim clearly *is* synthetic:

If, hypothetically, a coin had come up heads every time in one hundred blindly chosen tosses, then the coin would not be fair; more specifically, the propensity for Heads to come up would be greater than the propensity for Tails to come up._____(2)

It is clear that (2) is not analytic. Its antecedent clause tells us about one thing (the results of tosses of a coin), its consequent clause about something quite different (the existence of a propensity). It is both epistemically and metaphysically possible for (2) to be false. I take it as uncontroversial that, if (2) is true, it is synthetically true.

Now let us consider the relation between (1) and (2). Plainly, (1) is neither more nor less than the assertion that we have good reason to believe (2). The relation between (1, 2) is the same as the relation between:

We have good reason to believe P.________(3)

and

P.____________________________________(4).

We noted above that (2) is, very clearly, synthetically true. But, it will be argued, we must say we have at least some *a priori* reason to believe it. The argument for this conclusion is very straightforward: We have *a priori* reason to believe (1), and (1) tells us we have some reason to believe (2). Therefore we must have at least *some a priori* reason to believe (2). Since (2) is synthetic, we must therefore have at least some *a priori* reason to believe this synthetic proposition. We seem to be confronted here with an example of a synthetic proposition for which we can have some *a priori* reason.

How can this be? It may perhaps seem that we have obtained "something from nothing". We started out with an *a priori*, plausibly analytic proposition, and appear to have somehow derived a synthetic one. Surely there must be some kind of sleight-of-hand going on here?

The key move is from (1) to (2). And this move is of the same form as the move from (3) to (4). For simplicity, we will focus our discussion on the move from (3) to (4). In what follows we will assume that (3) is analytically true. (This assumption is not essential, but without it the argument becomes somewhat more cumbersome.) Now, if (3) is analytic, then we may presumably assert:

The degree to which it is rational to believe (3) is (very close to) one.__________(5).

That is:

The degree to which it is rational to believe: "We have *good* reason to believe P" is (very close to) one.________________________________(6)

Let us make the further assumption we can assign some number to the degree of reasonableness of P. Assume that we may rationally believe P to degree n, where n is some number greater than zero but less than one. Then (6) becomes:

The degree to which it is rational to assert: "We have reason to believe P to degree n" is one.____(7)

It has been argued above that (7) gives us reason to believe P. But we have not in any sense "got something from nothing" in doing this *provided that the degree of confidence with which we assert* P *is no more than* n. More generally, the move from:

"We have reason to believe P to degree n"

to

"P"

is permissible, so long as we make the degree of confidence we have in P "appropriately lower" than the degree of confidence warranted by "We have reason to believe P". More specifically, the move would seem to be permissible provided that our degree of confidence in P is no greater than *n*.

The move from (3) is a *contentfully* ampliative inference, but it need not be an *epistemologically* ampliative inference. The content of "P" is not wholly contained within the content of "There is good reason to believe P": there is a clear sense in which "P" says more than does "There is good reason to believe P". So, the inference is contentfully ampliative. But it need not be *epistemologically ampliative*. Suppose speaker A_1 knows it is rational to believe P with degree of confidence n, while speaker A_2 does in fact believe P with a degree of confidence not exceeding n. There seems to be clear sense in which A_2's *knowledge* does not exceed that of A_1. Therefore, the inference is not epistemologically ampliative. The inference need not be epistemologically ampliative if the conclusion P is asserted or believed with the appropriate degree of epistemic modesty.

If this is granted, then we can perhaps see a way in which Humean scepticism might be avoided. The Humean sceptic says we cannot rationally justify inferences that are (contentfully) ampliative, that is, where the conclusion has content not contained within the premise. But the argument just given, if sound, shows that is not so. An inference that takes us from a premise of the form "We have reason to believe P to degree n" to the conclusion "P" might have just this characteristic. In the example that has been discussed in this section, it has been argued that the premise "We have reason to believe P to degree n" can be analytic even though "P" is clearly synthetic. So, the inference is undeniably contentfully ampliative. But provided we assert P with no more than confidence of degree n, it intuitively seems to be a rationally justified inference. And so, it seems we do have here a rationally justifiable inference of just the kind denied by the Humean sceptic. We have perhaps found a chink in the Humean armour.

2.7 Is This Acceptable to a Moderate Empiricist?

It will be argued that the reasoning given above ought to be acceptable to the Moderate Empiricist. Let E be the observations that, *prima facie*, support some claim P. For example, E might be the observation that a hundred blindly selected tosses of a coin were all heads, and P is the claim the coin is not fair. We can represent the argument as follows:

If E, then P has epistemic probability n_______(8)

(8), we are assuming, is analytically true. So, (8) will be acceptable to a Moderate Empiricist. But we may also assume that we have established by empirical observation:

E____________________________________(9)

Since it is established by observation, E will also be acceptable to a Moderate Empiricist.

From (8) and (9) it follows by *modus ponens*:

P has epistemic probability n.______________(10)

Since (8), (9) and *modus ponens* are all acceptable to a Moderate Empiricist, so is (10). If E counts as knowledge – and there is no reason within Moderate Empiricism why it could not – then it seems as though (10) could, for the Moderate Empiricist, also count as an item of knowledge. Here we will assume that for the Moderate Empiricist it *does* count as an item of knowledge.

Now, it has been argued that we can, with certain provisos, infer from (10):

P____________________________________(11)

A person who believes P with degree of confidence n believes no more than that which they are licensed to believe by their knowledge of (10). So, given that the Moderate Empiricist can accept (10) as knowledge, so can the Moderate Empiricist accept P, *provided it is asserted with a degree of confidence no greater than that licenced by* (10). If P is accepted with appropriate epistemological modesty, it is a synthetic *a priori* claim that ought to be acceptable to a Moderate Empiricist.

2.8 A Consideration of Some Objections

It might perhaps be suggested that where the above argument goes wrong is right at its beginning. More specifically, it might be asserted that we actually *do not* have any *a priori* reason to believe (1): the claim that tells us if a hundred blindly chosen tosses are all heads then we have reason to believe the coin is not fair. Of course, it would be extremely implausible to hold that we have *no reason at all* to believe (1). But perhaps our reason for believing (1) is *entirely* empirical.

In considering this suggestion, it is worth reminding ourselves of a standard definition of *a prioricity*. To say we know P to be true *a priori* need not imply that we (can) know P to be true without any input from experience whatsoever. Rather, on a standard view, P is knowable (or reasonable) *a priori* if the only experience necessary to know it (or have reason to believe it) is the experience necessary to understand the meanings of the terms in it. And since sometimes some experience is necessary to understand the meanings of terms, sometimes experience is *required* for us to have *a priori* reason for a proposition. Very plausibly, some experience is needed to understand the meanings of many of the terms in (1). But now, suppose a person had all the experience necessary to understand the meanings of the terms in (1), and no more experience supportive of (1). Would such a person have at least *some* reason to believe (1)? I think our strong intuition is that they would have at least *some* reason. But, if this is granted, it follows by the argument given that such a person must have at least some reason to believe synthetic proposition (2): they would have some *a priori* reason to believe (1), and hence at least some, albeit possibly weaker, reason to believe (2). We can, it seems, have *a priori* reason for some synthetic claims.

Still, the feeling may persist that it is simply not possible to have *a priori* reason, of any degree of strength, for any synthetic proposition. It might further be suggested that it is the *contingency* of such propositions that makes it impossible for us to have any kind of *a priori* reason for them. If a proposition is contingent, it is true of *this* world, but *not* true in at least some other possible worlds. But: if we have not observed the actual world in which we live, how can we possibly have reason to believe it is one of the worlds in which P is true?

One possible response to this objection is the following: We cannot, perhaps, know *a priori* the features of the actual world, but we can perhaps have *a priori* reason to believe that certain worlds are more common than others. For example, we may have *a priori* reason to believe that worlds in which a blindly chosen coin but fair coin comes up heads a hundred times in a row are *rare* worlds. And if we have *a priori* reason to believe such worlds are rare, we can have *a priori* reason to believe we are probably not in such a world.

Perhaps the main source of resistance to the idea that we could have (good) *a priori* reason for believing something synthetic and contingent about the actual world is that the *source* of such belief would seem to be, as we have already noted, "deeply mysterious". How could we have good reason to believe anything about the world without perceiving it in some way or other? Perhaps a *partial* answer to this question is suggested by the above paragraph. We often make claims about "possible worlds". It will be assumed here that sometimes (at least) we do have good reason for the claims we make about possible worlds. But, however it is we come to have good reason to believe things about other possible worlds, it is *not* "by observation". Now, as noted above, sometimes we have reason to believe some possible worlds are rarer, or more common, than others. And this may give us probabilistic reason to believe something about our own world. How have we come to this probabilistic belief about our own world? The answer is: via that faculty that gives us reason to believe things about possible worlds, together with *a priori* probabilistic reasoning about the frequency or rarity of some worlds compared to others. Of course, this does not tell us *how* we come to have good reason for things about possible worlds. But: a philosopher who accepts that we can have reason to believe things about possible worlds, including that some are rarer than others, ought perhaps also be prepared to accept we can have *a priori* but probabilistic reason to believe some contingent and synthetic propositions.

2.9 Induction

If we are prepared to allow that there can be *a priori* reason for a synthetic proposition, an argument for induction can be developed, and follows fairly naturally from some considerations given in the previous sections.

We have, given the argument of the previous sections, *a priori* but defeasible reason for saying:

If a large number of crows have been observed, in blindly chosen locations, and they have all been black, then the propensity for crows to be black is greater than the propensity for them to be non-black.__________________________(12)

The case for saying there is synthetic *a priori* reason for (12) is plainly as good as that for saying we have *a priori* reason for asserting (as we did with (2)) that a particular tossed coin has a greater propensity to come up heads.

To say there is a greater propensity for crows to be black than there is for them to be non-black is very different from saying “All crows are black”. However, it will be argued, a natural extension of the reasoning that would lead us to (12) *can* be used to justify induction.

Suppose, again, we have observed a large number of crows in blindly chosen locations, and they have all been black. For definiteness, let us assume that we have observed all the crows in the city of Geelong (where Geelong as the location for our observations was chosen “blindly”) and all crows proved to be black. One possibility is that all crows everywhere, both in Geelong and outside it, are black. Another possibility is that while all the crows in Geelong are black, the crows outside Geelong are all non-black. But if the choice of Geelong as the place to observe crows was a “blind” choice – that is, made in ignorance of the colour of the crows there – then an improbable event would have occurred: the blindly chosen location of our observations would have happened to have coincided with the location of the all the black crows. Since this is surely unlikely, we have reason to believe it has not occurred. That is, we have reason to believe it is not the case that the crows outside Geelong are all non-black.

The “core” of the above reasoning can be represented as follows:

(A) If Geelong were an island of black crows in a sea of non-black crows, then a highly improbable event would have occurred: the blindly chosen location for our observations would have happened to have coincided with the island of black crows in a sea of non-black crows.

Therefore:

(B) It is probably not the case that Geelong is an island of black crows in a sea of non-black crows.

The move from (A) to (B), at least on the face of it, seems quite uncontroversial: It is of the form: If P then Q, probably not Q, so probably not P. This type of inference is standardly used, for example, in statistics. The crucial premise, therefore, would seem to be (A).

It is undeniably the case that (A) is very plausible. But the relevant question is: Do we have *a priori* reason to believe it? And the answer to this question surely *seems* to be "Yes": in the absence of any empirical information other than that necessary understand (A), a speaker surely, I think we are intuitively inclined to say, has good reason to believe it.

But can we do better than this appeal to intuition? Perhaps we can. Let us suppose that the number of crows in the world is M, and the number of crows in Geelong is N. Then, if it is assumed the probability that we should blindly choose any one crow to observe is the same as the probability we should blindly choose any other, the chances that we should have blindly chosen all and only the black crows (i.e. the crows in Geelong) to observe is given by the following expression:

$$\frac{N}{M} \times \frac{N-1}{M-1} \times \frac{N-2}{M-2} \times \ldots\ldots \times \frac{N-(N-1)}{M-(N-1)} \qquad (P)$$

Pretty clearly, if we assign N and M plausible numbers, the value of (P) will be very low indeed.[18] In fact, the only case in which the value of (P) is not less than ½ is that in which there are only two crows in the world, with the one black one being in Geelong and the non-black one being outside Geelong. And since, on the face of it, is seems *a priori* unlikely that the world should be exactly that way, we may surely conclude that it is *a priori* unlikely that we should have blindly chosen to observe an island of black crows in a sea of non-black crows.[19]

But does the argument just given actually establish that we have *a priori* reason for (A)? There are, unfortunately, a number of objections that can be raised.

2.10 The Principle of Indifference

One objection is that the above argument implicitly relies on the principle of indifference; and there are well known difficulties with that principle.

[18] By "plausible numbers" it is meant "numbers of crows that are in fact realistically likely to be in Geelong and in the rest of the world. So, there are presumably at least hundreds of crows in Geelong and at least millions elsewhere. Note also that this assumption does not in any question-begging way involve making an assumption that will assure us of our desired conclusion. The conclusion argued for here could, perhaps, be qualified to something like: "On the assumption that that there are "sufficiently many" crows in the world, it is rational to prefer enumerative induction." Also, "sufficiently many" does not here have to mean "*very* many". As the subsequent passage of the main text points out, the claim of the argument will still be true provided there is at least one crow in Geelong and one crow elsewhere.

[19] Perhaps strictly speaking the considerations just given only support the thesis that it is *a priori* highly likely that it is unlikely that we should have blindly chosen this.

Let us begin by stating the principle of indifference:

If a number of different states of affairs P1, P2, …., Pn are all possible, and we are not in possession of any reason to regard any of them as more or less likely than any other, then each one of P1, ….., Pn ought to be assigned the same epistemic probability._______________(PI).[20]

Now, the argument given above pretty clearly does rely on PI. The first term in the expression is: N/M. There are, we assumed, M crows in the world and N black ones, all of which are located in Geelong. In our first observation, we blindly choose one crow to observe. We choose a black one. Very plausibly, the chances of us choosing a black one is given by N/M since there are M crows in the world and N ways we could choose a black one. But note: this assumes that the choice of any one crow to observe is neither more nor less likely than any other. That is, the principle of indifference has here been assumed.

However, the principle of indifference has a well-known problem. The problem arises from the fact that applying the principle to two properties of the same set of objects, where the magnitudes of those properties are related non-linearly, can lead to contradictions. Although the properties with which we have here been concerned ("crowness" and "blackness") would not seem to be related in this way, we surely wish to be able to apply induction to *all* properties, and so we need to consider this difficulty for the principle of indifference.

One example of this problem for the principle has been described by B. van Fraassen.[21] Van Fraassen asks us to imagine a factory that produces squares of metal. There are, let us assume, two facts we know about the squares of metal.

(i) The sides of the length of the squares vary from 1 foot in length to 2 feet in length.
(ii) The areas of the squares of metal vary from one square foot to four square feet.

Plainly, (i) and (ii) would seem to be pretty much logically equivalent. But we seem to be led to a contradiction if we apply, in a natural way, the principle of indifference to (i) and (ii).

Suppose we are told there is a square in the next room that has been produced by the factory. What is it reasonable to believe about its dimensions? Let us assume we focus on (i), that is, that the squares vary in side-length from one foot to two feet. Applying the principle of indifference to (i), it seems natural to say there is a probability of ½ that the length of the square will be somewhere between one foot and 18 in., and a probability of ½ that it will be between 18 in. and two feet. Plainly, if

[20] The expression "The Principle of Indifference" is due to J. M. Keynes *A Treatise of Probability*, (MacMillan and Co. 1921), especially Chapter IV "The Principle of Indifference".

There are, of course, some refinements that can be made to the principle, for example, that the sum of the probabilities of P_1, …, P_n cannot be greater than one, and if P_1, …, P_n exhaust all the possibilities, the sum must be exactly one. But as these refinements do not concern the issues with which we are here concerned, we will ignore them.

[21] See van Fraassen *Laws and Symmetry* (Oxford Clarendon Press, 1989), esp. pp.307–309.

the length of a square is 18 in., then its area will be two and one-quarter square feet. And so we are naturally led to say that probably half the squares will have an area between one square foot and two and a quarter square feet, while the other half will have an area between two and a quarter square feet and four square feet. However, if we apply the principle of indifference to (ii), we get a different result! (ii) tells us the area of the squares varies from one square foot to four square feet. Applying the principle of indifference to this claim, we are naturally led to say that there is a probability of ½ that a square will have an area between one square foot and two and a half square feet, and a probability of ½ it will have an area between two and a half square feet and four square feet.

In summary, applying the principle of indifference to (i) leads us to say there is a 50% chance a square will have an area between one square foot and two and a quarter square feet, while applying it to (ii) leads us to say there is the same chance the area of a square will lie between one square foot and two and a half square feet. And this is despite the fact that (i) and (ii) would seem to say the same thing. Consequently, it has been widely accepted that there seems to be something wrong with the principle of indifference. And if the principle of indifference is wrong, it seems we are not entitled to assert that the probability of a blindly chosen crow (on our first observation) being black is N/M.

So, in summary, in this situation applying the principle of indifference, in a natural way, has led to contradiction. One possible response to this might be to conclude that the principle of indifference is irredeemably flawed. Another response might be to try to discover or work out ways of applying the principle that do not lead to contradiction. A number of ways of doing this have been explored.[22] But this approach would appear to leave us with the residual problem of working out which such way is best.

Here, however, a different strategy will be adopted. As noted above, a common way of dealing with the problem is to find *some* application of the principle of indifference that does not lead to contradiction. Here, however, it will here be argued that *on all possible applications* of the principle of indifference, the argument for induction advocated here will still go through. And so, whatever way of dealing with the puzzle might be adopted, the induction used here will be available.

Let us again assume the world contains M crows, and all and only the N black crows are in Geelong. Let us also assign to M and N reasonably "realistic" numbers. N is presumably at least in the hundreds, and M very much larger. We blindly choose a crow to observe. On what is presumably the most natural application of the principle of indifference, no one crow is more likely than any other to be chosen to be observed. So: the most natural application of the principle of indifference would lead us to say the probability of any one specific crow being chosen is 1/M. But we can imagine other possible applications of the principle of indifference. Let us call the specific crow chosen "Boris". Conceivably, a speaker might say: "For any crow,

[22] For an overview, see Hájek, Alan, "Interpretations of Probability", *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2012/entries/probability-interpret/>. Especially section 3.1.

it is either identical to Boris or it is not. So: on one possible application of the principle of indifference the probability of a crow identical to Boris being selected is ½ and the probability of a crow not identical to Boris being selected is also ½. Suppose we blindly choose Boris to observe, and he proves to be black.

Now, we have noted that we are here assuming that there is a plausible number of crows in Geelong – presumably in at least the hundreds. There is, we may assume, *more than one* crow in Geelong. And, of course, there are still crows outside Geelong. What is the probability that the next crow we blindly choose to observe will be one of the (black) Geelong crows? It is surely incompatible with anything that could be called "the principle of indifference" that the probability of us choosing a black Geelong crow is *one*. The principle of indifference tells us that if there are two or more possible events and we have no reason to regard one of them as more likely than another we ought to assign the same probability to them. The crucial point is that, in the situation envisaged, there are clearly *many* ways in which we could be said to have no more reason to believe one possibility rather than another. Suppose there are a million crows in the world, and a hundred in Geelong. Then, on what is perhaps the most natural application of the principle of indifference, the chances of us choosing one of the black Geelong crows is a hundred in a million, that is, one in ten thousand. But, of course, there might be other applications of the principle. We may say, for any crow, it is either a black Geelong crow or it is not. And this might suggest that an application of the principle on which the chances that the next crow we pick will be a black Geelong crow is ½. *But there does not seem to be any application of the principle of indifference on which the chances of us selecting a black Geelong crow is one.* And so long as the chances of us selecting a black crow are less than one, the probability that the first two crows we select for observation will be both be black Geelong crows must be less than ½.

In summary, the main point being made here is as follows. It is true that the principle of indifference can be applied in a number of different ways in any situation. And it is also true that in estimating the probability of us blindly choosing only the black crows in a sea of non-black crows we are using some form or other of the principle of indifference. But, it has been argued, on *all* possible applications of the principle of indifference, the probability of us blindly choosing to observe N crows and of them all turning out to be black becomes less than ½ as soon as N becomes greater than one. So, the fact that the principle of indifference can be applied in a range of incompatible ways to the same state of affairs hardly gives us good reason to doubt (A).

2.11 Objection: Other Inductive Inferences Can Be Made from the Data

However, of course, we are still some way from having reason for "All crows are black". It has been argued that if Geelong were an island of black crows in a sea of non-black crows and we blindly happened to choose that island as the location for our observations, a highly improbable event would have occurred. This does seem to give us good reason to suppose (given our observations) that:

All crows in Geelong are black while those elsewhere are non-black.____(13)

is probably false. But even if (13) is probably false, it still does not follow that we have reason to believe all crows are black. Perhaps the crows in Geelong are mostly black while half those outside are black, the other half white. Perhaps those outside Geelong are mostly black, but there is the occasional non-black one. Or perhaps all of the crows outside Geelong are black except for a single white one, say, in Stockholm zoo. And even this would be incompatible with "All crows are black". There are many possibilities other than "All crows are black" compatible with falsity of (13).[23]

However, it is not too difficult to see that the type of argument used against (13) can also be used against these other possibilities. One possible state of affairs compatible with our observations is:

All crows in Geelong are black, while those outside Geelong are 50% black and 50% white._______________________________(14)

But again, it is clear that if (14) were true, our blindly chosen location for observing crows would have happened to have coincided with an island of all black crows in a sea of crows that are both black and white. It is also clear that the chances of this happening are very slim. And so we have reason to believe (14) is false.

Of course, there are some possible states of affairs in which the blindly chosen location of our observations does not *improbably* coincide with the location of the black crows. Geelong is in the southern hemisphere. Perhaps all the crows in the southern hemisphere are black while all those in the northern hemisphere are non-black. Our chances of blindly choosing a location in which the crows are all black would, under these circumstances, presumably be about ½. While this is not "improbable", it is clear that it is lower than the chances of us blindly choosing a location in which crows are black in a world in which *all crows everywhere* are black. If all crows everywhere are black then the chances of us blindly choosing such a location will, of course, be one.

[23] This objection is made by J. Meixner and G. Fuller against BonJour's justification of induction. See Mexiner and Fuller "BonJour's *A Priori* Justification of Induction" in *Pre-Proceedings of the 26th International Wittgenstein Symposium*. S. Kostenbauer (ed) (2008)., pp.227–2

What this means is that if we say crows in the northern hemisphere are non-black, we are claiming something less likely will have occurred than if we say all crows are black. And this clearly gives us reason to prefer the hypothesis that all crows are black.

The same point can be made if even just one crow outside Geelong (say a white crow in Stockholm zoo) is non-black. In that case, a less-than-maximally-probable event would have occurred. We blindly chose our location for observing crows, and this location turned out to have only black crows. Even if there is only one non-black in the universe, the probability of our blindly chosen location for observing crows turning out to have all black crows is less than one. Only if all crows everywhere are black is it maximally probable that our blindly chosen location for observing crows should turn out to have only black crows.

It might be protested that the notion of justification used here is excessively weak. It was stated above that even if there is only one non-black crow in the universe, the probability of our blindly chosen location for observing crows turning out to have all black crows is less than one. But it might be pointed out in response that if the number of crows in the universe is very large, the probability of this is only marginally less than one. And if so, it might be protested, we are almost as justified in believing that that there is at least one non-black crow as we are in believing that all crows are black.

While this is so, there are two points to be noted. The first is that we still have *probabilistic* reason to prefer the conclusion that all crows are black to the hypothesis that there is at least one non-black crow. So, induction is on probabilistic grounds preferable to all alternatives to induction. And a number of sceptics about induction have denied this. For example, Karl Popper has asserted that the probability of an inductively arrived at generalisation never rises above zero.[24] There are also some defenders of induction who would claim to have done something less than show induction to be probabilistically preferable to alternatives. For example, Hans Reichenbach only claimed to have shown that induction will work if any method will work.[25] And F. J. Clendinnen took at his task in justifying induction to provide some grounds for preferring induction to its alternatives.[26] The sense of "justification" used here is something that some sceptics have asserted cannot be given and seems to be at least as strong as that has been given by some friends of induction.

The second point is that even if it is true that this justification of induction is rather weak, it does not follow that the case for *scientific realism* to be developed here is also correspondingly weak. A central notion used in the defence of realism given here is the Eddington inference. Although the justification of Eddington

---

[24] See, for example, Karl Popper *The Logic of Scientific Discovery*, p.364.

[25] Hans Reichenbach "The Pragmatic Vindication of Induction" reprinted in T. J. McGrew, Marc Alspector-Kelly and Fritz Alhoff *The Philosophy of Science*: *An Historical Anthology* (Willey-Blackwell, 2009), p.366–371.

[26] F. J. Clendinnen "Rational Expectation and Simplicity" in *What*? *Where*? *When*? *Why*? *Australasian Studies in the History and Philosophy of Science* edited by Robert McLaughlin (Springer, Dordrecht, 1982), pp.1–25.

inferences, to be given in Chap. 5, is logically like the justification given in this chapter to induction, it is also, plausibly, *rather stronger* than the justification given for induction.

The conclusion of an Eddington inference is not a universal generalisation but rather an existence claim. Let us assume all the crows we have examined have been black and have been at a range of locations, all within a 10-km radius of the centre of Geelong. An inductive inference from this data would be the universal generalisation "All crows are black". This makes a claim about a potentially indefinitely large number of crows. It would be falsified by the existence of just one non-black crow. But an Eddington inference from the same data might be "There exists at least one crow more than ten kilometres from the centre of Geelong." The existence of just one such crow is always sufficient to make this true, not matter how many crows there are in the universe. Provided such a crow outside Geelong exists, the conclusion of the Eddington inference would not falsified by any number of crows, whether black nor non-black, outside Geelong. The conclusions of Eddington inferences are much more logically modest than those of inductive inferences. Consequently, Eddington inferences are more likely to lead us from true premises to true conclusions than are inductive inferences from the same data.

This strengthens the case for scientific realism to be developed here. To repeat: even if it true that the sense of "justification" claimed here for induction is rather weak, this need not undermine the case for realism. Eddington inferences play a crucial role in the defence of realism, and Eddington inferences seem to be rather stronger than inductive inferences. We return to these themes in Chap. 5.

2.12 Another Objection: The Possible Influence of the Observer

A difficulty frequently raised objection against any attempt to justify induction appeals to the possible influence of an observer. All crows we have observed have been black, but perhaps it is our presence as observers that *causes* any non-black crows to *turn* black. However, it is not too difficult to see a similar type of reply to the above can also be given to this suggestion.

First, let us see how the idea we as observers might be causing crows to go black might seem to undermine (A). Recall that (A) says that, given that we have observed all the crows in Geelong and they were all found to be black, if it were asserted that Geelong is an island of black crows in a sea of non-black crows, then a highly improbable event would have occurred: the blindly chosen location for our observations would have happened to have coincided with the island of black crows in a sea of non-black crows. But if it is our observing crows that *causes* them to go black, no improbable event has occurred. Our choice of location for observation is still blindly chosen, but we are only observing black crows, not because they are all objectively and independently of our observation *black*, but because our act of observing them causes them to go black.

Here it is useful to distinguish between two distinct claims:

(1) All the crows we have observed have been found to be black because *the crows we have observed* have a propensity to go black when observed; other crows do not have that propensity.
(2) All crows, everywhere and everywhen, have a propensity to go black when observed, but when unobserved are some other colour.

It is easy to see that (1), at least, can be shown to be unlikely on the view advocated here. It is *a priori* unlikely that the crows we have blindly chosen to observe should be just the crows that have a propensity to go black when observed.

The matters raised by (2) are little more complex. Theses of observer-dependence can, broadly speaking, be divided in to empirical theses and philosophical or metaphysical theses. Empirical theses say that the observer exerts some sort of causal influence on the world where this causal influence is part of the familiar causal order of things. An example might be: Whenever I walk past a particular dog it is barking, and it is barking because it can see I am looking at it. Metaphysical theses of observer dependence, on the other hand, assert that the observer influences the world via some mechanism that is not part of the familiar causal order. The metaphysics of Berkeley is presumably an example of this type of observer dependence, as is, perhaps, the kind of observer dependence that has been claimed under some interpretations of quantum theory.

Empirical claims of causal dependence present no special puzzle for the view advocated here. They can be tested by straightforward empirical methods. (I might for example test the hypothesis it is my presence that is causing the dog to bark by observing the dog when I know he cannot see me.) Metaphysical claims of observer dependence are, however, another matter. And I admit that the approach to induction developed in this chapter provides no way of showing such metaphysical theses of causal dependence to be less likely than common-sense realism. But still, it will be argued, given the overall aims of this book we can still have good reason to prefer common-sense realism to metaphysical claims of observer dependence.

The aim of the book is a defence of the claim that we have *good reasons* for scientific realism. But this of course raises the question: "*How* good must these reasons be to be for our purposes "good enough"?" One answer, and the stance taken in this book, is that our reasons for scientific realist claims are "good enough" if they are, more or less, as good as the reasons we have for believing in the existence of familiar objects. One argument in support of this position is as follows. There is a view, often attributed to G. E. Moore, that common-sense claims such as "This is a human hand" are more certain than the claims of philosophy. On such a view, if an assertion of philosophy conflicts with common-sense, it is the philosophical claim that ought to be rejected. Such a view is relevant to the dispute between realist and non-realist views in science. If the scientific realist claims can be given a justification more or less as good as that for common-sense realist claims, then we have good reason to prefer scientific realist claims. Given such an aim, it is not necessary to provide a justification for the common-sense realist claims themselves. And so, given such aims, it is not necessary to provide a justification for common-sense realism over metaphysical claims of observer dependence.

2.13 Grue-Bleen Type Predicates

Another familiar objection appeals to grue/bleen type predicates. We could, it seems, apply the argument given here to the conclusion "All crows are blite", where, for example, a crow is "blite" if and only if it is, say, black and observed in the Geelong or white and observed outside Geelong. But we surely do not wish to embrace the conclusion "All crows are blite".

A full discussion of the problems arising from grue/bleen would, of course, be very extensive. The author has discussed the matter more fully elsewhere.[27] Here a brief outline of a response will be given.

A way of replying to the objection from grue-bleen type predicates arises naturally from the argument for induction given here. If we are permitted in using, in *ad hoc* way, grue-type predicates we have concocted in response to the data, then the justification of induction used here stalls at the very beginning. Suppose we blindly chose Geelong as the location for our observations, observed all the crows in Geelong and found them to be black. After having made these observations, we constructed the predicate "blite" and offered the hypothesis "All crows are blite".

If we arrived at the hypothesis "All crows are blite" in this way, then the starting point in our argument for induction cannot be established. For our argument to rule out a hypothesis such as "All crows in Geelong are black while those outside it are non-black", it must be the case that is unlikely that the location of the black crows coincided with the blindly chosen location of our observations. But if we *post hoc* define "blite" *after* having observed black crows in Geelong, it is plain no improbable event would occurred in us finding positive instances of "All crows are blite". And so we are unable to proceed any further with our inference to the conclusion "All crows are blite".

2.14 Concluding Remarks

The aim of this chapter has been to defend the claim it is possible to justify induction. It has had two more specific aims: (1) To defend the idea that it is possible to have *a priori* albeit defeasible reason for synthetic propositions. (2) To argue that it is possible to give a justification for induction, construed narrowly as enumerative induction. The defence of enumerative induction is probabilistic, in the sense that it

[27] See *Explaining the Success of Science* (Acumen, 2014), esp. pp.70–81.

has been argued that we have good reason to believe enumerative induction confers upon its conclusions a higher probability than other forms of inference.

This conclusion will be used in later chapters. One type of inference that places a key role in establishing Scientific Realism is the "Eddington inference". It will be argued in Chap. 5 that since we have probabilistic reason to prefer the conclusions of enumerative induction, we also have probabilistic reason to prefer the conclusions of Eddington inferences. The argument also underlies the defence of the notion of the independence of theory from data sketched in Chap. 6, and the use made of the Cosmological Principle in Chap. 8.

Chapter 3
The Skeptical Arguments – 2

In this chapter we continue the task of defending Scientific Realism against threats and challenges. In particular, in this chapter we consider:

- The Pessimistic Meta-Induction on the History of Science.
- The argument from the underdetermination of theory by data.
- The problem of "equivalent descriptions".
- An argument from Bayes' Theorem purporting to show the probability of a theory must always be zero.
- The argument from the "Experimentalists Regress".
- The argument from the allegedly unscientific character of realism.
- The argument from the theory laden-ness of observation.
- The objection from unconceived possibilities.

3.1 The Pessimistic Meta-induction on the History of Science

Perhaps the challenge to Scientific Realism that has received most discussion in recent years is the "pessimistic meta-induction on the history of science". A full discussion of this argument might easily take up an entire book: We will not attempt such a full discussion here. Instead, we will focus on the handful of cases from the history of science that are perhaps seen as constituting the most serious threat to Scientific Realism.

It is useful to start by explaining, in advance, what will be argued in the following sections. As we have noted, there seem to be two dimensions to Scientific Realism: the existence dimension and the behaviour dimension. There are a small number of "key cases" from the history of science that seem to most strongly support the pessimistic meta-induction, and so create a difficulty for scientific realism. However, it will be argued that these key cases present more of a difficulty for the *existence* dimension of realism than for its *behaviour* dimension. If this is so, then

J. Wright, *An Epistemic Foundation for Scientific Realism*, Synthese Library 402, https://doi.org/10.1007/978-3-030-02218-1_3

what is required to strengthen the realist's position is some kind of good reason to accept the *existence claims* made by scientific realists. If, in at least some cases, we have good, independent reason for accepting existence claims about unobervables, then perhaps we need not be worried by the examples from the history of science that seem to cast doubt on existence claims that have been made by realists. It is this good, independent reason for the truth of existence claims which, it will be argued, is supplied by Eddington inferences.

Let us now turn our attention to the cases that seem to give support to the pessimistic meta-induction. Prominent amongst these cases are: the phlogiston theory of combustion, the caloric theory of heat and the theory of the lumeniferous ether. Another case, perhaps less well known, is that of Rankine's thermodynamics.

3.1.1 The Phlogiston Theory of Combustion

On the face of it, if any theory creates a difficulty for realism it is the theory of phlogiston.[1] It was highly explanatorily successful: It was, around 1780, able to explain as much as the rival oxidation theory of Lavoisier.[2] It was, arguably, as simple as oxygen theory.[3] It had even enjoyed novel predictive success.[4] And yet, it turned out it was wrong in its explanatory claims: it is not the case that phlogiston is causally responsible for the phenomena of combustion. And it is also wrong in its central existential claim: there is no such thing as "phlogiston". Phlogiston theory therefore seems on the face of it to be highly likely to create a difficulty for realism: it was a very successful theory that did not even come close to the truth.

[1] The theory of phlogiston is given as an example creating a difficulty for realism by Larry Laudan in his "A Confutation of Convergent Realism" in Jarrett Leplin *Scientific Realism* (University of California Press, 1984), pp. 218–249, especially p. 231.

[2] See Alan Musgrave "Why did oxygen supplant phlogiston?" in *Method and Appraisal in the Physical Sciences*, edited by Colin Howson, Cambridge University Press, 1976, pp.181–210.

[3] See Musgrave, *op cit*.

[4] For the advocates of phlogiston theory the combustion of some substance C was not oxygen combining with C, but phlogiston being given off by C. This way of viewing things led Joseph Priestley to offer his own interpretation of what happened when mercury oxide (what he called "the precipitate *per se*) was heated to produce pure mercury. Since mercury oxide was, in his view, dephlogisticated mercury, the process of heating mercury oxide to produce pure mercury must have been the process of dephlogisticated mercury re-absorbing phlogiston. And so the resultant "air" produced by this process, he reasoned, must be air from which the phlogiston had been removed. He referred to this as dephlogisticated air. So, for Priestley, what we now call "oxygen" was "dephlogisticated air". But now, for Priestley, in combustion phlogiston is given off by burning substances: it stops once the air has become saturated with phlogiston and cannot absorb any more. On Priestley's view, therefore, "dephlogisticated air" (oxygen) ought to have a capacity for supporting combustion greater than that of ordinary air: since it is *de*phlogisticated air it ought to have more "room" to absorb phlogiston than ordinary air. And this was observed to be the case – things were found to burn more easily and energetically in dephlogisticated air. Thus phlogiston theory led to a novel prediction that was subsequently found to be correct.

Although phlogiston theory is undeniably wrong,[5] on closer examination it turns out to present less difficulty for realism than might at first be thought. Recall again the two dimensions of Realism: the truth dimension and the existence dimension. For "full-blown" realism about phlogiston theory to be correct, two conditions must hold: Condition (1): phlogiston theory must make (more or less) true claims about what is going on in combustion and Condition (2): the entity (ies) postulated by phlogiston theory must exist. Of course, phlogiston does not exist, so phlogiston theory does not meet Condition (2). But, it will be argued, phlogiston theory comes rather closer to making true claims about what goes on in combustion than initial impressions might suggest. Phlogiston theory, it will be argued, comes quite close to meeting Condition (1).

It is useful to start by giving a broad outline of the main features of the theory. Phlogiston theory said that combustion was the emission of a substance called "phlogiston". When something burned, it was not combining with oxygen, rather, it was expelling phlogiston. What we now regard as oxidation was seen as the process of "dephlogistication". Processes such as the burning of wood, the heating of a metal in air to produce a calx, and also the respiration of animals, were seen as processes of dephlogistication. And processes that we now call "reduction", or the expulsion of oxygen, were seen as the acquisition of phlogiston. As a number of commentators have remarked, phlogiston can be seen as a kind of "anti-oxygen".[6] In something the same way that an electric current can be represented as either a stream of negatively charged particles flowing one direction or a stream of positively charged particles flowing in the opposite direction, so many chemical reactions could be represented as oxygen flowing in one direction, or phlogiston flowing in the opposite direction. Phlogiston theory gets things wrong, but there does seem to be a sense in which it gives us a *mirror image* – a *reversed* mirror image – of what is actually going on.

As James Ladyman has noted, there appears to be a sense in which phlogiston theory is "isomorphic with" the correct account.[7] We could go a fair way in turning it in to the correct account simply by replacing "dephlogistication" with "oxidation"

---

[5] One way in which phlogiston is wrong is that it misrepresents what is going on when something burns. We know now that when something burns it combines with oxygen: burning is a form of (chemical) combination. But phlogiston theory says that in combustion something (phlogiston) is given off from the burning substance. For phlogiston theory, burning is not combination but expulsion or separation.

There is a sense in which a substance that burns *does* lose at least something in combustion: the electrons in its outer shell. But I take it that no one would seriously suggest this vindicates phlogiston theory. Presumably advocates of phlogiston theory conceived of phlogiston as a kind of substance, on a par with substances such air or water or carbon. A quantity of electrons is presumably too dissimilar from what advocates of phlogiston theory had in mind to count as a candidate for what they were getting at.

[6] This is discussed in James Ladyman "Structural Realism versus Standard Scientific Realism: The Case of Phlogiston and Dephlogisticated Air", *Synthese* 2011, volume 180, pp.87–101. The idea that phlogiston could be viewed as "anti-oxygen" is also raised in Mikhail Volkenstein *Entropy and Information*, Birkhauser Physics (2009), p.6.

[7] See Ladyman, *op cit*.

and "phlogistication" with "reduction". And at least some of the theoretical claims that follow from phlogiston theory are straightforwardly *true*. For example, phlogiston theory tells us that the process that occurs when mercury oxide is heated is *the reverse* of the process that occurs in combustion and respiration.

The fact that phlogiston theory gets the structure of the causal processes right, but not their direction, could be seen as supportive of a form of Structural Realism.[8] It gets on to some aspects of the truth about theoretical processes even though labouring under one central, false assumption.

Phlogiston theory is perhaps the theory that first comes to mind when we think of an example of a successful theory from the history of science that was "completely wide of the mark": but on closer inspection it seems it actually *isn't* very wide of the mark.

But, of course, the fact still remains that phlogiston theory is not true *in toto*. There is no such thing as phlogiston. Realism about phlogiston is wrong because it fails to satisfy the "existence dimension" of realism. Yet, it was explanatorily highly successful. This suggests that the inference from "T is a successful or good theory" to the conclusion "The unobservable entities postulated by T *exist*" is not as reliable as some realists might wish.

3.1.2 The Caloric Theory of Heat

Another theory that, on the face of it, would appear to support the pessimistic meta-induction is the caloric theory of heat. This theory has turned out to be false. But, it was for a while highly successful empirically. In this section it will be argued that the relation between the caloric theory of heat and pessimism is like the relation between phlogiston and pessimism. This case does not, or does not clearly, support pessimism with respect to our claims about the laws things obey, but does support a degree of pessimism with respect to the existence dimension of realism.

According to the caloric theory, heat is a substance, more specifically, a particular sort of gas. This gas was held to be so penetrating or permeating it could pass through tiny pores that were believed to exist even in apparently solid substances such as metal. Crucially, various parts of a quantity of caloric were held to be mutually repulsive. From these simple assumptions about the nature of caloric, many of the observed properties of heat could be derived. Heat can be transmitted through solid substances. Since caloric was a gas, putting more heat in air (or any other gas) caused it to expand. And, given appropriate additional assumptions, this provided a natural explanation of Charles' Law that, provided pressure remains constant, the volume of a gas is proportional to its temperature. Since quantities of caloric are mutually repulsive, heat tends to spread out from hotter places to cooler places and become evenly distributed throughout an object. Also, since heat is, on this view, a form of matter and matter is conserved, it follows that the total quantity of heat in

[8] This is explored in Ladyman, *op cit*.

the universe must also be conserved. From these results, a number of laws of thermodynamics follow. The principles of the "Carnot cycle" of engines were also derived from these aspects of the caloric theory of heat.[9] The caloric theory was therefore a highly successful theory. But, of course, it was false: heat is not a substance of any kind: caloric *does not exist*. Since it was a successful theory that turned out to be false and whose central explanatory entity turned out to not exist, the caloric theory of heat would *prima facie* be a positive instance of the pessimistic meta-induction.

However, as has been pointed out by Stathis Psillos, the situation is perhaps not that simple.[10] Briefly: although the caloric theory of heat is wrong about heat being a type of substance, it is right about very many other things. More specifically, Psillos argues the theory is right, or very nearly right, about the laws of experimental calorimetry, adiabatic change and Carnot's theory of the motive power of heat. These consequences of caloric theory were, according to Psillos, independent of the idea that heat was a kind of substance. The example of phlogiston, Psillos concludes, does not support pessimism.

Psillos' conclusions have, however, been contested by Hasok Chang.[11] Chang argues that caloric theory was highly successfully empirically even though it was *not* even close to the truth about what was going on at the theoretical level. Chang sees this as constituting a difficulty for Scientific Realism and supportive of the more sceptical view of Laudan.

It is worth examining Chang's argument in a little more detail. Chang refers to the version of caloric theory developed by Laplace. According to Laplace's theory, heat was a gas, more specifically, it was a gas consisting of point-like particles. We will refer to these as the particles of caloric. These particles had two key properties: they repelled other particles of caloric, and were attracted to particles of ordinary (non-caloric) matter. These properties were, according to Chang, essentially used by Laplace in deriving the empirically successful predictions of caloric theory. But, says Chang, subsequent scientific research has found no evidence whatsoever that anything like the particles postulated by Laplace actually exist. So, Chang concludes, caloric is an example of an empirically successful theory that that does not seem to get things even approximately right at the theoretical level. It therefore constitutes a difficulty for realism.

However, it seems to the present author that Chang somewhat overestimates the extent to which caloric is a difficult case for realism. Scientific Realism *would* perhaps be confronted with a puzzle or embarrassment if we had here a case of an empirically successful theory that did not even come close to getting things right at the theoretical level. But – while it is certainly true that the *particles* of caloric

---

[9] "Reflections on the motive power of fire" by S. Carnot (Paris, Bachelier, 1824).

[10] See S. Psillos "A philosophical Study of the transition from the caloric theory of heat to thermodynamics: Resisting the pessimistic meta-induction, *Studies in the History and Philosophy of Science*, 25 (1994), pp.159–190. For an opposing point of view, see Hasok Chang.

[11] See Hasok Chang "Preservative Realism and Its Discontents: Revisiting Caloric" in *Philosophy of Science* vol. 70 (2003), pp.902–912.

postulated by Laplace do not exist – it is not entirely fair, I think, to suggest that caloric theory gets things completely wrong at the theoretical level. We have just considered phlogiston theory. There is certainly no such thing as phlogiston, but still phlogiston theory is not completely in error as a description of what is going on at the theoretical level. As we have noted, we could get an account that is, from our point of view, pretty close to the truth simply by replacing "oxygen" with "phlogiston" and reversing the direction in which the phlogiston is held to be moving. And, it will be suggested, a comparably simple transformation could turn caloric theory in to something fairly close to the truth.

One thing that is wrong with caloric theory is that it makes what we might nowadays call a "category-mistake": it says that heat is a substance where now we would say it was a property of substance – more specifically, it is a form of motion. The caloric theory of Laplace said a unit of heat was a tiny particle; we would say a unit of heat was a unit of kinetic energy. And if we replace "particle of caloric", as it appears in Laplace's theory, with something like "unit of kinetic energy", we do get something that looks rather like the account we believe now.[12] Laplace thought that particles of caloric repelled each other. But now, let us consider the particles that make up a gas. The motion of each particle is random with respect to the motion of each other particle. So, if the particles are not constrained within a vessel, they will (eventually) tend to move away from each other. Also, the greater the kinetic energy of each particle, the more rapidly it probably will be moving away from other particles. It will look "as if" the units of kinetic energy are being repelled by other units of kinetic energy. Of course, we do not now think that there is any repulsive force between the units of kinetic energy. In a gas unconstrained by the walls of a vessel, they behave (more or less) as if there are repelled by each other simply because their motion is random with respect to each other. But, seen in this way, the picture of Laplace recognisably resembles our own picture: it is not an entirely different picture altogether.

Laplace also thought that particles of caloric were attracted to particles of ordinary matter. This can also be translated in to something that closely resembles the account we now believe. Laplace's "law" that particles of caloric were attracted to ordinary matter can be re-expressed as the "law" that particles or units of heat were attracted to quantities of matter lacking in heat. And if we interpret "particle of caloric" as "unit of kinetic energy", we again get a picture recognisably similar to our own. If a small collection A of particles with relatively high average kinetic energy interact with a much larger collection of particles B with comparatively less average kinetic energy, the average kinetic energy of the particles in A will become closer to the average kinetic energy of those in B. We can say, roughly, that regions of higher average kinetic energy tend to spread their kinetic energy throughout regions with lower kinetic energy. So, regions with concentrations of high kinetic energy behave as if they are "attracted to" regions of lower kinetic energy. From this

[12] Perhaps it is more correct to say that "particle of caloric" is to be replaced by "quantity of mass possessing kinetic energy equivalent to work capable of being done by a particle of caloric". But as the latter expression is somewhat cumbersome, I will simply speak of "unit of kinetic energy".

perspective, our view of what is going on at the theoretical level does broadly resemble Laplace's.

On the view just sketched, the account that the caloric theory gives of what is going on at the theoretical level bears some similarity to ours. If we replace "particle of caloric" with "unit of kinetic energy", we can get a picture that can be recognised as a distorted, rough sketch of our own, rather than as an entirely different picture altogether. The case of caloric presents less of a challenge to realism than Chang's remarks would seem to suggest.

Still, of course, there is one respect in which caloric theory is flatly wrong. Heat is not a substance. Caloric theory was – plausibly – a good explanation of the phenomenon associated with heat, even though it got it wrong concerning *what exists* at the theoretical level. As with phlogiston, the caloric theory of heat is wrong in its central existence claim.

Again, we see that a case from the history of science gives some support to pessimism about *existence* claims. The reliability of the inference from "T provides a good explanation of some phenomena" to "The unobservable entities postulated by T *exist*" would seem to be cast in to doubt.

3.1.3 The Theory of the "Lumeniferous Ether"

Another case that seems to support pessimism is the version of the wave theory of light advanced by Augustin Fresnel. One consequence of Fresnel's theory is that there ought to be a white spot in the middle of a perfectly round shadow. This consequence of the theory was derived by Poisson, who regarded it as a *reductio ad absurdum* of the theory. However, when observations were performed by Arago, it was found that the white spot was in fact there, exactly as the theory predicted.[13]

One reason why this case is of special interest is because the white spot would seem to be very clearly a *novel* prediction of the theory. Fresnel's theory was therefore a theory that had novel predictive success. And yet, as Laudan has pointed out, Fresnel's theory was false.[14] According to Fresnel's theory, light was a type of wave in the ether. But since there is no such thing as the ether, Fresnel's theory was *wrong*.

Of course, a natural response to this is to say, although Fresnel's theory was, strictly speaking false, it was not false through-and-through. Fresnel, we can surely say, was wrong about the ether, but at least partially right about the nature of the wave character he attributed to light.[15] More specifically he was right in his idea that

[13] A summary can be found in Eugene Hecht, *Optics* (Pearson Education Limited, 2014), pp.496–497.

[14] Laudan *op cit*, p.225.

[15] More specifically, Fresnel's theory is based on the idea – originally due to Christiaan Huygens – that each point in a wave of light acts as the source or origin of another, spherical wave of light emanating out from that point. This is known as "Huygens' Principle". The propagation of a wave

each point at the front of a wave of light itself acts as a source from which an expanding, spherical, wave of light emerges.[16] Saying Fresnel was right about some, although not all, of the wave characteristics of light might naturally be seen to support a form of Structural Realism. But it also casts doubt on the reliability of inference to the best explanation, where the entity inferred *to exist* is not observable.

Let us briefly remind ourselves of just why it was that, up until the early twentieth century, people believed in the ether. There seemed to be strong evidence that light had a wave character. The confirmation of Fresnel's theory was evidence of this, as were phenomena such as diffraction. But, it seemed, if there is a wave, there must be some underlying medium that is doing the "waving". For ocean waves, the medium is water, for sound waves, it is the air or any other form of matter through which the sound is transmitted. So, it seems very plausibly to follow, there must be some medium through which light waves move. But it was also known that light can pass through a vacuum: a bell jar does not grow dark once air has been removed from it. Light can also cross the airless space from the stars to the Earth. So, it was thought, there must be, filling all of space, some undetectable substance that acts as the medium for waves of light.

To a person in the nineteenth century, the hypothesis of the ether must surely have seemed to be, not just a possible or plausible explanation of how light can travel through a vacuum, but the *only possible* explanation. And so, by default, it must have been the best available explanation. And yet, we now regard it as false. So, we see, again, that inference to the existence of an unobservable entity on the grounds that it supplies us with an explanation (in this case, surely the best *available* explanation) of some phenomena has not fared well in the subsequent history of science.

In summary, Fresnel's theory is, at least according to what we currently believe, true in part but also false in part. It seems to be right about certain aspects of the wave character of light. So, it does not in any unqualified or unrestricted way support the pessimistic meta-induction. But it seems to be wrong in what it has to say about the *existence* of the entity by which light waves are propagated. Once again, using IBE to postulate the existence of an unobservable entity has turned out to be wrong.

---

of light is the result of the "adding together" of all these spherical waves. (More precisely, it is the adding together of these waves in the direction in which the beam of light is travelling.) Huygen's Principle is accepted in modern physics. See, for example, M. Born and E. Wolf, *Principles of Optics*: *electromagnetic theory of propagation*, *interference and diffraction of light* (Cambridge University Press, 1999), p.986.

[16] Born and Wolf, *loc cit*.

3.1.4 Rankine's Thermodynamics

In the mid-nineteenth century the Scottish physicist W. J. M. Rankine developed a theory of the nature of heat.[17] Rankine accepted that heat was a form of motion. But, to say merely that heat "is a form of motion" leaves it open just what it is that is in motion. Rankine offered a theory of what is "in motion" that is, by our lights, not remotely close to the truth, and even bizarre. And yet, it was surprisingly empirically successful. As Keith Hutchison has argued, this case would appear to give support to a pessimistic meta-induction.[18]

We can begin by giving a broad sketch of Rankine's theory of the nature of heat. Rankine accepted that heat was a form of motion, but as noted above, it was not for him the mean kinetic motion of atoms or molecules moving randomly with respect to other atoms or molecules. It was rather a form of motion taking place *within* the atom. Rankine conceived of the atom as a nucleus surrounded by an approximately spherical "atmosphere". It was for Rankine a certain type of motion within this atmosphere that constituted, and was responsible for the effects of, heat. The atmosphere of an atom was on Rankine's view made up of numerous tiny, tapered cylinders. The narrow end of each tapered cylinder was located on the surface of the nucleus of the atom. The larger end of each tapered cylinder coincided with the outer boundary of the atom. Each tapered cylinder took up the same space as a set of radii emanating from the centre of the atom (or, rather, the surface of the nucleus) and terminating at the boundary of the atom. Inside each of these tapered cylinders was a fluid. The fluid was thought to be rapidly rotating around the radius that lay along the centre of each tapered cylinder. And it was the rotation of this fluid that was, for Rankine, responsible for the phenomena of heat.

Briefly, Rankine thought that the more rapidly the fluid rotated, the greater the force it would exert against the walls of its containing cylinder. This in turn would cause the tapering cylinders to exert pressure outwards, away from the nucleus. The outer boundary of an atom would thereby exert more force on the outer boundaries of other, surrounding atoms. In this way an increase in the motion of fluid in the cylinders gave rise to the phenomenon we know as pressure.

Rankine derived a number of consequences from his theory. They are listed below:

(i) The equation of state for steam and imperfect gases.
(ii) The cooling experienced by carbon dioxide in the Joule-Thompson expansion experiments.
(iii) The entropy function.
(iv) That the specific heats of perfect gases will be constant.

---

[17] Rankine, W. J. M. "On the mechanical action of heat, especially in gases and vapours" in *Transactions of the Royal Society of Edinburgh*, **20**, pp.147–190. (1853).

[18] See Keith Hutchison "Miracle or Mystery?: Hypotheses and Predictions in Rankine's Thermodynamics" in S. Clarke and T. Lyons *Recent Themes in the Philosophy of Science* (Kluwer Academic Publishers, 2002), pp.91–120.

(v) That saturated steam will have negative specific heat.[19]

The case of Rankine might seem to be especially puzzling for the realist. Rankine's theory was empirically successful and, like the other examples so far considered, the unobservable entities it postulated turned out to not exist. But – unlike the other examples – the claims it makes about what is going on at the theoretical level do not seem, by even the most generous of standards, to be close to the truth. Rankine's is not recognisable as a distorted version of our own picture, it does instead seem to be an *entirely different picture*.

Of the cases considered here, this one perhaps presents the greatest difficulty for realism. But still, it will be argued, the realist need not be *too* embarrassed by it. In short, the reason why the realist need not be too worried is because, although Rankine's theory was empirically successful, it was not in all *other* respects a good theory.

Theories are underdetermined by the observations on which they are actually based. So, for any body of observations, there will be a number of theories that can explain the data. Even if some observations are novel, there will still be a number of ways of explaining the observations. Realists need not be committed to saying that there will only be one theory that can provide an explanation (whether a good explanation or a bad explanation) of some body of observations. But, if a realist wishes to be able to maintain that we can have *good reason* to say a theory in some domain is true, then the discovery of *several good*, but substantially different, theories in that domain might not be welcome. If there were a number of theories "tied for first place" as best, it would be difficult to adopt a realist view of that domain: *Which* theory ought we to be realist about? A realist would rightly be worried if there were in a domain a number of theories "tied for first place".

But it is difficult to see Rankine's theory as being one "tied for first place" as an account of the phenomena associated with heat. One feature of theories that is widely accepted as making them good is simplicity, and Rankine's theory is not very simple. Compared to the theory that heat is mean kinetic energy of molecules, or even to Laplace's theory, Rankine's mechanism of numerous tapering cylinders surrounding atomic nuclei and filled with rapidly rotating fluid seems very elaborate. Newton once said: "Nature is pleased with simplicity, and affects not the pomp of superfluous causes".[20] And if Nature *had* adopted Rankine's mechanism, then Nature *would*, it seems, have adopted highly "superfluous causes" given that the phenomena could, surely, have been caused in much simpler ways.[21] Rankine's

---

[19] The account of Rankine's theory given here is derived from Hutchison, *op cit*.

[20] This is a part of Newton's First "Rule of Reasoning in Natural Philosophy", from his *Mathematical Principles of Natural Philosophy*.

[21] One student of mine remarked that Rankine's mechanism had a "steampunk" quality about it. "Steampunk" is a style in fashion and art that imagines what our modern world might be like if we had continued to rely on the same general types of mechanical device as those prevalent in the ninetheenth century, such as the steam engine. So, an imagined "steampunk" world might have flying machines, computers, communications systems and so on powered by steam engines, clockwork etc. rather than jet-turbines and electronics. The mechanical devices imagined (or constructed) in "steampunk" can be of fantastical and grotesque complexity.

theory is empirically successful, but that does not mean it has *all* the properties that would make it a good theory.

These considerations show that the realist need not be too worried about the case of Rankine. The most any realist would surely want to say is that if a theory has (novel) predictive success *and* has the other properties (including simplicity) associated with theoretical goodness, then we thereby have good reason to believe the theory is in some sense close to the truth. It would be an embarrassment for the realist if there were a theory that had all the good-making features yet was not even close to the truth, but Rankine's theory is not an example of that sort. It is, by our lights, not even close to the truth; but neither does it seem to be in all respects a *good* theory. A realist need not be too troubled by it.

3.1.5 Summary of the Historical Cases

We have considered a number of examples from the history of science that have been claimed to present a difficulty for realism. It has been argued that these examples do not provide *unqualified* support for pessimism. But neither would they seem to support a fully optimistic realism. All the cases considered would seem, from our contemporary point of view, to be true in part but also false in part. But there is also a pattern that can be discerned in the way all the cases considered fell in to error. All fell in to error in postulating the *existence* of unobservable entities on the grounds that they provided (what was seen as) the best explanation of some observed phenomena. Phlogiston, caloric, ether and the tapering cylinders of Rankine turned out *not to exist*.

`But, of course, most of the theories we considered were not wholly false. The theory of the ether was wrong about what light waves were waves of, but right about aspects of the structure or behaviour of the waves. Phlogiston theory in a broad sense *mirrored* the causal processes taking place in combustion. Caloric theory could easily be transformed in to something recognisably similar to our theory. These theories were, in some sense, more or less close to the truth about what is going on at the theoretical level.

Our survey would seem to support the idea that we cannot reliably make the inference from the fact that a theory provides a good explanation to the conclusion that the unobservable entities it postulates exist. However, it is important to note that it does not follow from this that we cannot have a good reason *of any sort* for the claim that some class of unobservable entities exist. There may be some other way of rationally justifying the claim that the unobservable entities exist.

It is appropriate at this point to note that *not all existence claims concerning unobservable entities have had a bad track record*. There are *some* existence claims that have proved to be very resilient. Perhaps most obviously, the claims that molecules and atoms exist seem to have stood up well to testing. One possible response to this might be to say that inference to unobservables on the grounds that they furnish us with a good explanation of the phenomena does, *sometimes*, lead us to the truth, even if does not always do so. But this is not the only possible response. Another possibility is that there may be a route to unobservable entities *other than* inference to the best explanation. More specifically, it may be that there is some route to unobservables that more reliably leads to truth than IBE. This is the option to be explored and defended in this book. The notion of an "Eddington-inference", it will be argued, more reliably leads to true claims about unobservable entities than does inference to the best explanation. Moreover, it will be argued the existence of atoms and molecules can be established by means of Eddington inferences.

3.2 The Underdetermination of Theory by Data

An important source of doubt about the idea that science gives us truths comes from the thesis of the underdetermination of theory by data.

The expression "the underdetermination of theory by data" can refer to (at least) two quite distinct theses. One such thesis, largely associated with W. V. Quine, is that our theories are underdetermined by all *possible* evidence.[22] Another thesis is that the scientific theories we have actually accepted are underdetermined by the empirical evidence upon which they are actually based.

The first, Quinean, thesis is surely much more controversial than the second. As we acquire more empirical evidence, more potential explanations will be ruled out. There may or may not be more than one candidate left standing once all possible evidence is in.

The second thesis is much less controversial, and is undoubtedly true. There are historical examples of two theories both able to account for the empirical data at the time. For example, around 1550 the geocentric Ptolemaic and heliocentric Copernican systems could both account for the observed data equally well. But more generally, it is always possible to construct two or more alternative ways for explaining any finite, actually obtained body of data if we are prepared to embrace a complex, *ad hoc* system of hypotheses. Suppose, for example, that T is a theory that can explain some body of data D and T* is a theory that, while it can explain some parts of D, is apparently falsified by other parts of D. We can generally save T* from refutation by augmenting or modifying it in some *ad hoc* manner. We will refer to the modified version of T* as TA*. Then T and TA* will provide us with two, incompatible, ways of explaining the same data. Since it always seems possible to do this, actual scientific theories are in this way always underdetermined by the data on which they are based.

[22] Quine argues for this thesis in a number of places, including "On the reasons for the indeterminacy of translation" *Journal of Philosophy*, 67 (1970), pp.178–183, *Word and Object*, Cambridge, Mass., MIT press, 1960.

Both forms of the thesis can constitute a challenge to scientific realism. In this section we first consider the (non-Quinean) version of the thesis that says theories are underdertermined by the data on which they are actually based. We then consider the (Quinean) thesis they are undetermined by all possible data.

On the face of it, the fact that our theories are always underdetermined by the data on which they are based would appear to be a powerful reason for doubting that we have good reason to believe our theories. If, at any point in time, the data available in any domain can be explained by two or more theories T_1, T_2, … then the question arises: Can we be said to have good reason to believe that one of them (T_1, say) is true? Perhaps we cannot. And: if this is so, it would seem to follow that realism with respect to T_1, or to any other scientific theory, cannot be well founded.

The realist, however, does have available some replies to this sceptical argument from underdetermination. The premise of the sceptical argument is that there will generally be more than one theory that is capable of explaining any given body of data D. But this need not show, for example, that all the theories that explain D are equally credible. This only follows if we make the additional assumption that the *only* rational support for any theory comes from the fact that it explains D. But perhaps there can be other types of evidence for a theory. And, of course, one of the main theses defended in Chap. 2 is that there *can* be *a priori* reason to prefer one theory over another that explains the same actual phenomena. It was also argued that we have *a priori* reason to prefer enumerative inductions from some observations over other possible conclusions.

It might perhaps be protested that even if it is agreed that enumerative induction furnishes us with a way in which one theory can have more support than another that explains exactly the same data, this fact is of little help to the realist. Enumerative induction might make it more reasonable to believe observable but to date unobserved entities are relevantly similar to those that we have so far observed. But it would not appear to give us any reason to believe in the existence of unobservable entities, such as the micro-entities postulated by science. And so, it would seem unlikely to be of any help with the cases of underdetermination with which we are here concerned. However, one of the aims of this book is to argue that this is not exactly the case. The notion of an Eddington inference has a crucial role, on the view advanced here, in establishing cases of realism. But, it will be argued, Eddington inferences can be given a justification closely analogous to the justification of induction given in Chap. 2. The argument to be given says there is good *a priori* but defeasible reason for accepting inferences to the existence of entities not detectable by our observational techniques. If this is right, the underdetermination of theory by actual data need not constitute a good argument against realism.

It is perhaps worth emphasizing the importance of the *a prioricity* of the justifications of induction and of Eddington inferences. Suppose these justifications were not *a priori,* but empirical and wholly *a posteriori*. Then, our account of what makes a scientific theory good would, presumably, derive whatever credibility it had from the data it was able to explain. Let us call such a purely empirical account of what makes our theories good EE (for Empirical Epistemology). According to EE, the theories that have actually been accepted by scientists will, we may assume, be judged to be better than other possible theories that can explain the same data. For example, according to EE, the theory "All crows are black" may count as good while "All crows in Geelong are black, the rest are green" may count as bad. But, of course, if the only evidence in support of EE is actually obtained empirical evidence, and theory is underdetermined by such evidence, there will be other purely empirical accounts of what makes a theory good. There may be some other such account, which we will call EE*. Perhaps, according to EE*, the theories that are to be counted as good are those that say that after, say, the year 3000, the future will be unlike the past. Or EE* might say simple theories about events before 3000 are good, while for any claims about events after 3000, complex theories are to be preferred. And so, it seems to follow, if we only allow purely empirical evidence we would have no way of dealing with the objection raised against realism from the underdetermination of theory by actual data. *A priori* justifications are required.

Let us now consider the Quinean thesis of the underdertermination of theory by all possible data. This, too, would seem to pose a threat – although of a slightly different sort – to realism. According to this thesis, there could be two or more theories T and T* that both explain all *possible* observations. A (possibly) stronger version of the thesis would say that T and T* would be, in some substantial sense, genuinely *different* theories: neither would be any sense just a "notational variant" of the other. The existence of such pairs of theories would, possibly, be a potential embarrassment for realism. If T and T* explain all possible observations then they must be "about" the whole universe: they would be, in Quine's terminology, "systems of the world".[23] And so, there would seem to be a clear sense in which they must be about the same thing: the whole universe. But now, if they are genuinely different theories, and they are about the same thing, it would clearly seem to follow that they cannot both be true. So: either one of them, or both of them, must be false.[24]

But, this would seem to have the effect of making truth *inaccessible*. Suppose T was in reality true while T* was false: then no possible observation could tell us that T was true. And if both were false then no possible observation could reveal to us the true theory (ies). Underdetermination would have the effect of making truth inaccessible in principle. And so, it would seem to follow, we could never have any reason for any realist claim.

However, it is clear that this argument can be given the same reply as the previous one. It assumes that the only evidence that there can be for a theory is the observations it can explain. But it has been argued that this assumption is false.

We may also note that that the argument currently under consideration, relying as it does on the thesis of the underdetermination of theory by all possible data, would seem to be more open to question than the one previously considered. That

[23] W. V. Quine "On Empirically Equivalent Systems of the World" in *Erkenntnis* **9**, (1975) pp.313–328.

[24] One author who has consistently defended the thesis that Quine's thesis that theory is underdetermined by all possible data is Lars Bergstrom. See, for example, his "Underdetermination and Realism" *Erkenntnis*, 21, pp.349–365, 1984, and "Quine, underdetermination and skepticism" *Journal of Philosophy*, 90, pp.331–358, 1993.

our theories are underdetermined by the data on which they are actually based seems to be quite uncontroversial. That theories would be undetermined by all *possible* data seems much less clear.[25]

3.2.1 *Laudan and Leplin on Underdetermination*

An argument due to Larry Laudan and Jarret Leplin, and further developed by Leplin, aims to show that it is possible to deal with difficulties (apparently) raised by underdetermination, and the related phenomenon of empirical equivalence, *empirically* and without recourse to *a priori* justifications of simplicity and the like.[26] But it will here be argued that the approach they develop does not work. It will further be argued we do need to rely on *a priori* justifications of criteria for discriminating between good and bad theories.

Laudan and Leplin develop their argument by exploring the relations that exist between the thesis of the underdetermination of theory by data and the thesis that there exist pairs, or sets, of theories that have exactly the same empirical consequences, that is, the thesis of *empirical equivalence*.

Pretty clearly, if there were two empirically equivalent theories T_A and T_B, we could never have any purely empirical reason for preferring one to the other. And this might be thought to show that it could never be rational to prefer one to the other. And this in turn would seem to show we would never be rationally entitled to adopt a (fully) realist view of the domain that T_A and T_B were about. But Laudan and Leplin argue that actually there is no threat to rational theory choice (and hence to the rationality of realism) here. Briefly, they argue that if the underdetermination thesis were correct, we could never be justified in asserting that two (or more) theories really were empirically equivalent.

Their argument is as follows. Suppose theory T is about unobservable entities of some sort; say, electrons. If T is *only* about electrons then T by itself might have *no* observational implications at all. It will only have observational implications when considered in conjunction with other auxiliary theories and background beliefs that relate the behaviour of electrons to our apparatus, or to other things that can be observed. Call these "other theories and background beliefs" B. The conjunction

[25] It is worth noting that in what was perhaps his fullest discussion of the issue, Quine refrained from actually asserting the thesis the there will be more than one theory capable of explaining all possible data. See Quine *op cit*. Quine argues that for the undedetermination thesis to be true, there must be two or more *genuinely different* theories that can explain all the data. But this, of course, raises the question: "Under what conditions are two apparently different theories *genuinely* different rather than merely "notational variants" of each other?" Quine develops an account of the conditions under which two theories are to be regarded as "genuinely different". He then professes agnosticism on the question of whether or not there would be two or more such genuinely different theories capable of explaining all possible observations.

[26] See L. Laudan and J. Leplin "Empirical Equivalence and Underdetermination" in *Journal of Philosophy*, **88**, (1991), pp.449–472.

T&B, we will assume, *does* entail certain predictions about what we can observe. Assume T&B entails O.

Now, Laudan and Leplin note that if the thesis of the underdetermination of theory by data is *generally* correct, the theories and background beliefs that comprise B will themselves also be underdetermined. If so, there may be some *other* set of theories and background beliefs, B*, that can account for the same empirical data as does B. Moreover, since B* is *different* from B, it may be that T&B* entail *different* observational predictions from those predicted by T&B. We will say T&B* entail O*.

What this shows is that, for a theory about unobservables such as our theory T of electrons, there is no such thing as *the* empirical content of T. T will have one set of empirical consequences relative to B, another relative to B*, yet another relative to some third set of background theories B**, and so on.

Laudan and Leplin argue that this casts arguments for scepticism from empirical equivalence in a rather different light. Suppose T_A and T_B are two different theories of (for example) the behaviour of electrons. If, hypothetically, T_A and T_B *were* empirically equivalent despite being *different* theories, then it seems we could have no empirical grounds for preferring the claims of T_A to the different claims of T_B. But this obviously assumes *there is such a thing* as two different theories being empirically equivalent. The analysis offered by Laudan and Leplin undermines that assumption. There is, on their view, no such thing as empirical equivalence *simpliciter*: there is only empirical equivalence relative to one or another set of additional background assumptions.

There is another, closely related consideration that would appear to further undermine any argument against realism from empirical equivalence. Suppose T_A and T_B are empirically equivalent according to *the currently accepted* background beliefs B. Then: it is possible that further empirical findings might lead us to replace B with B*, and that relative to B*, T_A and T_B are inequivalent. It would then become empirically possible to choose between the two theories T_A and T_B. That two theories are *currently* empirically equivalent need not mean they will forever be so. The most that the currently obtaining empirical equivalence of T_A and T_B would mean is that we are not currently able to empirically tell whether we ought to adopt a realist view of T_A rather than T_B. But future developments might yet reveal how we could tell. And this is something with which the realist could presumably live.

However, it will be argued it is far from clear just how effective the argument of Laudan and Leplin actually is. Most obviously: their argument would merely seem to show that two currently empirically indistinguishable theories T_A and T_B *might* become empirically distinguishable in the future. But the argument presented gives us no assurance they *will* become empirically distinguishable.

There is another, perhaps more serious difficulty with their argument. Suppose B are the background beliefs currently accepted, and that relative to B, T_A and T_B are empirically equivalent. We will assume that B* is the set of background beliefs that will become available at some point in the future, and that relative to B*, T_A and T_B are *not* equivalent. However, it does not follow from the fact that *T_A and T_B* are not equivalent relative to B* that there will be *no* pairs of theories that are equivalent relative to B*. Perhaps, with some *ad hoc* tinkering, T_A can be turned in to T_A*, where the latter is empirically equivalent to T_B relative to B*. Presumably, this could always be done if we were liberal enough in the degree of *ad hocery* we are prepared to allow. One case of empirical equivalence might go, only for another to pop up.

Leplin is aware of this objection and gives a number of replies.[27] He objects that empirical equivalence obtained as a result of liberal enough *ad hocery* is "trivial", and a science that permitted it would be incoherent and non-explanatory. However, it is not entirely clear how an attribution of "triviality" constitutes an *objection*. Of course, it may well be *trivially easy* to turn T_A in to T_A*, where the latter is empirically equivalent to T_B if we are entirely liberal in our use of *ad hoc* hypotheses. But the fact that it is trivially *easy* to come up with T_A* would not explain how we might *have good reason to believe* T_B rather than T_A*. We return to the significance of this point shortly.

Leplin also says that if science exhibited unrestrained liberality in its use of *ad hoc* background hypotheses then it would be incoherent and non-explanatory. But it is not clear how either accusation necessarily follows. Suppose "coherence" is taken to be "logical consistency". On the face of it, it would seem pretty likely that we could still easily enough get empirical equivalence if the only constraint on the background hypotheses we were allowed to use was logical consistency. If "coherence" is interpreted more strictly – as something like "fitting together naturally and plausibly" – then it may be harder to secure empirical equivalence. But: we have yet to see why theories that are "coherent" in this stricter sense ought to be seen as more worthy of rational acceptance than those that are not.

Leplin claims that if science were so liberal and unconstrained in its use of background hypotheses that empirical equivalence could always be obtained, such science would cease to be *explanatory*. But to the present author it is not entirely clear how this follows. It is surely one thing to say highly *ad hoc* explanations are widely thought to be *bad*, but another entirely to say they are not really explanations at all. And of course if we are to say that such background hypotheses are to be ruled out on the grounds that they *are* highly *ad hoc*, complex, and so on, we are then confronted with the question: What makes such hypotheses unworthy of rational acceptance?

Leplin suggests we ought to respond to the fact that it is trivially easy to get empirical equivalence if we are completely unconstrained in our use of background hypotheses by placing an *epistemic constraint* on what background hypotheses are permissible. It seems to me that Leplin's reasoning here is perhaps as follows. The empirical content attributed to T will depend upon which background assumptions are adopted. If we adopt background assumptions B, then the empirical content of T will be, we may say, CT_1; if we adopt background assumptions B*, the empirical content of T will be CT_2, and so on. So, we may have a number of possible candidates for the empirical content of T. These candidates will be CT_1, CT_2 and so on. Which, of any, of these candidates is to be preferred? A natural suggestion is that the

---

[27] See Jarret Leplin *A Novel Argument For Scientific Realism* (Oxford University Press, 1997), esp. 153–157.

one that is to be preferred is the one that is derived from T *together with the most strongly independently confirmed set of background beliefs.* More generally, “hypotheses” about the content of T deserve to be taken seriously only to the extent that the background beliefs used in generating them have some degree of independent confirmation.

In summary, Leplin’s position appears to be as follows:

(i) Claims of empirical equivalence that follow “trivially” from the unrestrained use of *ad hoc* background hypotheses are not to be taken seriously.
(ii) The only attributions of *empirical content* to scientific theories that do deserve to be taken seriously are those that follow from (the theories themselves when conjoined with) independently confirmed background hypotheses.

However, it is clear (ii), by itself, is not sufficient to ensure we will never have any examples of empirical *equivalence*. To get the latter conclusion, we need to add to (ii) some assumption like: “If we only use independently confirmed background hypotheses, two different theories about unobservables, such as our theories T_A and T_B about electrons, will not come out as empirically equivalent.” But as Andrei Kukla has noted, this assumption need not be true.[28]

This might lead us to an investigation of how likely or unlikely it might be for T_A and T_B to be found empirically equivalent if we only used independently confirmed background hypotheses. Since, however, this question might seem rather difficult to answer, here a different line of thought will be developed.

Let us begin by looking a little more closely at some of the main features any such background hypothesis B must possess if it is to count as well confirmed. Again, assume our T is a theory of electrons, and T by itself has no empirical consequences, but does so when conjoined with B. So: what must B be like if, when conjoined with T, it gives “empirical import” to the claims of the otherwise empirically inaccessible claims of T? Suppose T tells us that under certain theoretical conditions C a particle with such-and-such theoretical, unobservable properties P will exist. This claim of T’s in itself has, let us assume, no observational consequences. If B is to give it empirical consequences, then B must say, or entail, something like “If there is a particle with such-and-such theoretical, unobservable properties P, then observations O will be obtained.” That is, B must itself make claims about the same types of theoretical, non-observational entities and states of affairs as does T itself.

Now, if there is *independent confirmation* of B, then there will, presumably, be some observations O it explains. But, given that B is itself of a high degree of theoreticity (similar to that of T itself), there will, surely also be highly *ad hoc*, complex theories B*, B** that are, relative to some *other* background hypotheses, empirically equivalent to B. On what grounds, then, are we to prefer B to its empirically equivalent rivals?

[28] See Andrei Kukla “Empirical Equivalence and Undertermination”, *Analysis*, **53**, (1993), pp.1–7.

A natural response, from within Leplin's general framework, might be to say that we are only entitled to assume B, B*, B** and so on *will* be empirically equivalent if we allow unrestrained liberality in our choice of what might be termed the "background background" auxiliary hypotheses that confer empirical content on B, B*, B** and so on. But if we only use, say, *independently confirmed* "background background" hypothesis C, then perhaps B, B*, B** and so on will turn out to be inequivalent and empirically distinguishable.

But: it follows from what has been said that C will be of a degree of theoreticity comparable to T, B, B* and so on. Therefore, we will need some way of being able to empirically distinguish between the good hypothesis C from other, bad, background hypotheses C*, C** and so on. Perhaps this can be done with auxiliary hypothesis D…. But an infinite regress clearly threatens.

How then might we be able to discriminate between the *good*, independently confirmed background hypothesis B that explains some observations E, and the bad *but* empirically equivalent ones B*, B**and so on that we could come up with if we are permitted unrestrained liberality in our way of constructing hypotheses? It has just been argued that the enterprise of trying to do so by constructing "background background" hypotheses that would enable us to empirically distinguish B, B*, B** and so on, leads to an infinite regress. How, then, is it to be done?

Unless we are to accept scepticism, it seems to the present author that the foregoing discussion points towards the need for non-empirical criteria for distinguishing good hypotheses from bad. If attempting to use empirical criteria leads to an infinite regress, what alternative is there? And the approach of using non-empirical criteria, such as the *a priori* justification of inductions, of Eddington inferences and some other criteria, is the approach used here.

3.2.2 Stanford on Realism and Underdetermination

An influential discussion of the relation between Scientific Realism and underdetermination has been given by P. Kyle Stanford.[29] Briefly, Stanford argues that existing defences of the underdetermination thesis fail to undermine realism. But, he says, if the attack on realism is supplemented by what he refers to as the argument from unconceived possibilities, then realism *is* seriously undermined.

Stanford argues for four points:

(i) The mere assertion of the *logical possibility* of theories, just as good as our current best theories but fundamental different from them, does not by itself seriously endanger realism. For realism to be *seriously* endangered, something more than the *mere logical possibility* of such alternatives to our currently accepted best theories is required.

---

[29] See P. Kyle Stanford *Exceeding Our Grasp: Science, History and the Problem of Unconceived Alternatives* (Oxford University Press, 2006), esp. pp. 9–17.

(ii) There are techniques for "algorithmically" generating alternatives to at least some of the descriptions of the world given to us by physics, but again, these do not seriously challenge realism. Consider, for example, a description of a Newtonian universe such that the centre of mass of the universe is at rest. There are, clearly, other permissible descriptions of such a universe which do not say it is at rest but is rather moving in a straight line at a constant velocity. There are indefinitely many such descriptions because there are indefinitely many directions in which the universe could be moving and indefinitely many constant velocities. Presumably, no possible observation could give us reason to prefer any one of these descriptions to any other. But examples of this sort, Stanford maintains, surely do not seriously undermine realism. A realist could, for example, very plausibly only claim to be a realist about for example Newton's laws but not about the speed or direction of motion of the universe as a whole. More generally, algorithmically generated alternatives would not seem to endanger the core of what the Realist wants to say.

(iii) There are other techniques for generating empirically indistinguishable alternatives to our current theory. For example, perhaps we get the observational and experimental results that we in fact do, not because molecules, electrons and so on are actually real, but because we are living in a simulation designed to make it seem to us that they are real. While much recent discussion would seem to accept that something along these lines is a serious possibility, the doubts it raises are surely not peculiar to the philosophy of science and to Scientific Realism. Scenarios of this sort, Stanford points out, raise *general* epistemological difficulties: They do not just raise doubts about the reality of electrons, but also of tables, chairs and so on.

(iv) Stanford acknowledges that there are some serious examples of underdetermination. He cites as examples of non-trivially different but plausibly empirically equivalent theories: Newtonian physics with gravity as a force within flat space-time and a Newtonian system that replaces gravity as a force with curved space-time, Special Relativity and Lorentzian mechanics, and Bohmian hidden-variable quantum mechanics and Von-Neumann-Dirac formulations of quantum mechanics. But he points out that these examples are rather few in number and hard won. The existence of such a small number of cases does not, he asserts, constitute a serious or wide-ranging challenge to realism.

Stanford's more general conclusion is that the thesis of underdetermination, by itself, does seriously endanger realism.

Is what Stanford says right? It is, at least to the present author, difficult to find anything objectionable about Stanford's point (ii): It is a plausible thesis that "algorithmically" generated variants of existing theories do not constitute a very great threat to the realist. The other points, however, do raise a variety of issues.

Let us begin by considering point (iii). Suppose an underdetermination thesis is defended on the grounds that we might, for example, be living in a simulation. Then the idea that electrons are real would be cast in to doubt, but equally so would the reality of tables and chairs. Under such a scenario, electrons and tables *might* be epistemically more or less on a par.

Now, it seems very plausible to say that at least *part* of what the Scientific Realist aims to do is to show that the case for the existence of electrons is more or less as good as the case for the existence of tables and chairs. On the view adopted here, this is a part of what the realist needs to do. But now, we have just noted that defending the underdetermination thesis via, for example a simulation hypothesis would seem to lead to the conclusion that a belief in electrons *is just as good as* the belief in tables and chairs. However, of course, this would not be a *defence of realism*. If we are living in a simulation then the belief in the reality of tables and a belief in the reality of electrons are *both wrong*. What a realist wants to be able to do is show it is *rational to believe* in the existence of both tables and electrons.

How, then, ought we to respond to arguments for underdetermination from the idea that we might be living in a simulation? It seems to the present author that an acceptable response is the following. There is a philosophical tradition that would seem to keep the rationality of our belief in tables and chairs intact, while leaving it as open question whether it is similarly rational to believe in the existence of things like electrons, quarks and so on. An early statement of this perspective can be found in Aristotle's *Ethics*. Aristotle wrote: "What all men say is so, I say is so. The man who destroys this foundation for our belief is not likely to put in its place anything more plausible." However, perhaps its best-known expression was given by G. E. Moore when he claimed to prove the existence of the external world by holding up his hand and asserting that at least one human hand exists.[30] Moore has been widely interpreted as providing a general reply to the arguments of sceptics. A sceptic might provide us with an argument against, for example, the existence of material objects, or for some other contrary-to-common-sense thesis. To accept the sceptic's conclusion, we must accept both the truth of their premises and the validity of the inferential steps in their argument. Moore claimed that it is in general *not* rational to accept the conclusion of the sceptic's argument since it is more likely that there is something wrong with the sceptic's argument than it is that our belief in common-sense "Moorean facts" like "A human hand exists" should be wrong.

Here a Moorean view of matters will be accepted. On this view, it is rational to continue to believe observation sentences like "This is a table" or "The needle on the meter is pointing to "5"", even though there exist arguments which, if sound, would show *all* observation sentences about material objects to be false. Although these sceptical arguments *exist*, on the Moorean perspective it is more likely that the arguments are somehow flawed than it is that our observation statements are, at least typically or mostly, false. More specifically, on this Moorean perspective, arguments for underdertermination on the grounds that we might be living in a simulation, or deceived by Descartes' evil genius, are not sufficient to show we ought to reject common-sense realism about material objects. Common-sense realism remains intact.

[30] G. E. Moore "Proof of an External World", in *Proceedings of the British Academy*, 25, (1939).

But, of course, this Moorean perspective is not enough to establish the truth of *Scientific Realism*. While it is surely a Moorean fact that tables exist, it is not a Moorean fact that electrons do. *More work* needs to be done to establish the rational acceptability of Scientific Realism. And on the view advocated here this is what is done by Eddington inferences.

Let us now consider (i) and (iv). Do these succeed in showing that the underdetermination thesis does not pose a threat to realism? Or might it perhaps be the case that underdetermination poses a greater threat to realism than is suggested by (i) and (iv)?

Let us begin by considering (iv): that the existence of a *fairly small number* of known serious examples of underdetermination do not constitute a great danger to realism. Is this so? Stanford is presumably right in saying that that only a small number of serious examples have actually been discovered or constructed. But it is perhaps tempting to protest that it does not necessarily follow from this that there are only a small number of such cases "out there", waiting to be discovered. And: it might be pointed out that there some lines of thought that suggest that the number of possible alternative theories "out there" might be rather larger than those we have actually managed to construct. After all, Stanford himself says: the actually-discovered alternatives are "hard won". Constructing them has not been an easy or trivial task: it cannot be done "algorithmically". It is perhaps of a comparable degree of difficulty to constructing the first (known) explanation of some body of phenomena. And so, we have a possible explanation – quite independent of how many alternative explanations might be "out there" – of why not many alternatives have been produced: constructing them is *difficult*. Possibly the number of alternatives out there is very large but not many have been constructed *because* we find it difficult to come up with them.

The idea that there might be a lot more alternatives out there than we have managed to come up with is supported by another line of thought. In constructing an explanation, a scientist will of course, use concepts, or ideas, or notions. But it is surely very plausible to suppose that the concepts that the scientist uses will either be concepts with which the scientist is already familiar, or in some way constructed out them. It seems rather less likely that the scientist will formulate a theory out of entirely new or original concepts that have not been in some way constructed or derived from already known concepts. And if this is granted it does seem rather plausible that the alternative explanations we will come up with will constitute only a subset (possibly quite a small subset) of all possible explanations.

So, in summary, it seems as though there might be a stronger challenge to realism than seems to be allowed by (iv). Such a stronger challenge could be developed if a case could be made for saying that there really do exist a significantly large number of (explanatorily satisfactory) alternatives to the theories we have so far actually formulated. The idea that there might be significantly many of these as yet unconceived of ways of explaining phenomena is developed by Stanford himself in to his main argument against realism. The argument from unconceived possibilities is not an argument against realism wholly independent of the argument from underdetermination. It is rather a way of strengthening the argument from underdetermination.

It is a line of thought that helps to move beyond the bare assertion that alternative explanations are for all we know possible. It helps to block the inference from "only a handful of serious cases of underdetermination have actually been found" to the conclusion that "there aren't many serious cases of underdetermination "out there"".

But is Stanford's argument from unconceived possibilities a good one? The task of responding to this particular challenge to realism is undertaken in Sect. 3.8, below. It is argued that the notion of Eddington inferences does supply us with a way around the difficulties Stanford raises for the realist.

There is also a point that needs to be made in response to (i). In his discussion, Stanford seems to mean by a "theory" something similar to what working scientists might mean by a theory. A "theory", on such an interpretation, is not *merely* a collection of sentences: rather, it is, roughly, something that hangs together as a unified, explanatory whole. If a "theory" is construed in this sense, then perhaps we are *not* entitled to assume that there will be many such "theories" explaining any body of data.

However, we do not necessarily have to interpret the term "theory" to mean something like "unified explanatory whole". We might take "theory" to mean merely "collection of sentences", no matter how complex and un-unified it might be. Construed in this way, it surely is true that for any body of data there will be indefinitely many "theories" that explain the body of data.

But, of course, it is tempting to respond to this that very many of these possible theories will surely be very implausible candidates for truth. Perhaps, if "theory" is construed in this very liberal way, the underdetermination thesis will be undeniably true. But the truth of *this* version of underdetermination, it may be asserted, need not constitute any threat to realism.

However, let us consider exactly why this particular version of the underdetermination thesis might seem to not be a threat to realism. The suggestion was that if we allow sets of sentences, no matter how complex or un-unified, to count as "theories", then it does not seem very likely that such theories are likely to be true. But this evidently is based on the assumption that complexity, or a lack of unity or simplicity, is a sign of untruth. This assumption may, of course, be correct. But what grounds, if any, do we have for saying this?

The question with which we are now confronted is a version of the familiar philosophical question: What grounds, if any, do we have for regarding features of theories such as simplicity, or unity, as signs of truth? This question is considered in the next chapter, and there it is argued that as yet we do not have good grounds. But let us for the moment note the following point: if we assert that the (certainly true) form of the underdetermination thesis currently under consideration really does not pose a serious threat to realism, we would seem to be implicitly accepting that complexity is correlated with a reduced likelihood of truth, and it is not clear what, if anything, entitles us to make that assumption.

3.3 The Problem of Equivalent Descriptions

Related to the problem of underdetermination is the problem of "equivalent descriptions". In his paper "Realism About What?", Roger Jones points out that many theories, including classical mechanics, quantum mechanics and general relativity, can be given a number of equivalent formulations.[31] For example, classical mechanics can be formulated in terms of point masses interacting by means of action-at-a-distance or in terms of fields interacting by local causation. This raises a problem for the scientific realist. Suppose we wish to be realists about classical mechanics. What entities, specifically, are we to be realists about? Are we to say that the world consists of point-masses, or fields, or perhaps something else? More generally, if we are to be realists about our best theory T, whatever that may be, what are we to be realists about if T admits of a number of equivalent descriptions?[32]

One possible response to this difficulty is to move towards some form of "Structural Realism".[33] Another is to say that our ontological commitments need not be determined only by the nature of our best available theory, but also by coherence with metaphysical beliefs.[34]

However, the considerations of the previous section suggest another way of responding to the difficulty that Jones raises. Jones' argument assumes that we arrive at our ontological commitments *via* our best theory: we believe to be real whatever our *best theory* says to be real. And if our best theory will be able to be given a number of different formulations, then our ontological commitments will therefore be indeterminate. But, we need not accept that we must arrive at our ontological commitments via our best theory. There *may* be another route to realism. And if this is so, then *Jones* argument plainly fails. Moreover, that there is a route to realism other than via our best theory is one of the main themes of this book. Eddington inferences, it will be argued, provide another route to realism.

[31] See Roger Jones "Realism About What?" *Philosophy of Science*, 58, (2), pp.185–202 (1991).

[32] A number of authors have recognised that, in one way or another, the existence of equivalent descriptions constitute a challenge to realism. The idea is a theme in much of the work of Hilary Putnam in the 1970s and 1980s. See, for example, *Reason*, *Truth and History* (Cambridge: Cambridge University Press, 1981) and "Realism and Equivalence" in *Philosophical Papers vol* 3 *Realism and Reason*. It should be noted, however, that the object of Putnam's attack is perhaps "metaphysical realism" rather than "scientific realism". In his *What is this thing called science?*" Alan Chalmers argues that the existence of equivalent descriptions constitutes a problem for scientific realism.

[33] The *locus classicus* of Structural Realism is John Worrall "Structural Realism: the best of both worlds" in *Dialectica*, 43, pp.99–124 (1989). An application of Structural Realism to the specific case of different equivalent formulations of Newtonian mechanics can be found in John Wright "Realism and Equivalence" in *Erkenntnis*, July 1989, vol 31, pp.109–128.

[34] This option is defended in Alan Musgrave "Realism About What?" in *Philosophy of Science*, 59, (4), pp.691–697. (1992).

3.4 Bayes' Theorem and the Probability of Theories

A number of theorists, including Karl Popper, have argued that the probability of a scientific theory never rises above zero, no matter how many positive instances of it have been observed.[35] Evidently, to say that the probability of theory is always zero is at least in very great tension with the idea we have good reason to believe it. And it would seem to be in very great tension with the sort of claims the realist wishes to say can be backed by good reasons. In this section the argument will be criticised.

In the first chapters of this book, we so far have used – and will continue to use – a particular way of defending realism from attack. Many of the arguments against realism that we have considered are based on the assumption that we arrive at realist claims via our best theory: it is assumed that, if a realist claim R is to be justified it is to be justified by deriving it from our (current) best theory. And, one response to these arguments is that the route to realism used in this book does *not* go merely via our best theory. Now, it might perhaps initially be thought that we could use this strategy here. Popper has argued that the probability of theories is always zero. Popper's argument might constitute a compelling argument against realism if it is assumed that we arrive at our realist claims via our theory. But, if we arrive at our realist claims in some other way, then Popper's argument would seem to pose no threat to *our* form of realism. And since here we do not arrive at realism via our best theory, it might at first be thought Popper's argument fails to apply against the position advocated here.

Unfortunately, however, in this case this strategy will not quite do. Popper's argument, if good, does not merely apply against scientific theories that postulate unobservable entities: it applies against *all* unrestricted universal generalisations. And, of course, such generalisations are used in stating any scientific realist claim. On the view advocated here, scientific realist claims have a "behaviour" dimension. For example, to adopt a realist view of electrons is to say electrons exist *and behave more or less as our theory says they do*. But to say that electrons behave more or less as our theory says they do is to make a universal generalisation: it is to say that all electrons conform, at least approximately, to certain general laws. So, Popper's argument, if good, would invalidate the approach advocated here.

Popper's argument relies on Bayes' Theorem:

$$\Pr(T, E) = \frac{\Pr(T\&E)}{\Pr(E)} \qquad \text{\_\_\_\_\_\_\_\_\_\_(1)}$$

Where Pr(T, E) is the probability of theory T being true given evidence E, Pr(T&E) is the probability of T and E together being true, and Pr(E) is the probability of obtaining evidence E.

Popper argues that if T is a universal generalisation, Pr(T) must be zero and hence that Pr(T&E) must also be zero. He further argues that since Pr(E) must

[35] One formulation of this argument is given in Karl Popper *The Logic of Scientific Discovery*, p.364.

always take some value above zero, the RHS in (1) must be zero, and therefore the value of Pr(T, E) must always be zero, no matter what the nature of E.

Here it will be argued that Popper's argument fails. It will, very roughly, be argued that, plausibly, Pr(E) must *also* have a value of zero.[36] A more precise statement of the position to be defended here is that either Pr(E) has a value of zero, or else we are not entitled to make any assumption concerning whether it is zero or non-zero. In either case, Popper's argument fails to go through.

Let us begin by considering what the expression "Pr(E)" means. This expression refers to the *a priori* probability of us obtaining empirical evidence E.[37] To say this is the *a priori* probability of us obtaining E is to say it is the probability of us obtaining empirical evidence E prior to us obtaining any empirical evidence whatsoever. It will be argued here that there are good reasons for saying that the prior probability of obtaining any observations E, no matter what the specific nature of those observations, will always be zero.

Why might we be tempted to think that the value of "Pr(E)" must be *above* zero? Let us suppose that we are performing an experiment, hoping to measure the value of some property. Our apparatus is connected to some measuring device. The device has a face with markers numbered one to ten, and a needle that will, when the experiment is performed, point to one, but only one, of these markers. We then perform the experiment and the needle, let us assume, points to marker seven. Prior to us knowing to what marker the needle will point, the probability that it will point to seven is, it is surely natural to assume, 0.1. We may repeat the experiment many times, and the larger our body of data, the lower will be the probability of us actually obtaining that specific body of data. But, provided our data is only ever finitely large, the probability of us obtaining that specific body of data will always, surely, remain above zero.

Here the thesis will be defended that the above argument goes wrong right at the beginning. There is a plausible argument showing that the prior probability of even just one observation (say, the needle pointing to marker number seven) is zero.

We have our apparatus and connected to it is the measuring device. We then perform the experiment. But: how do we know the only possible effect of the experiment will be that the needle on the measuring device will point to one of ten

[36] It should be noted that there are other ways of replying to Popper's argument. One influential approach to the problem was developed by Colin Howson. See Howson's "Must the logical probability of a theory be zero?" in *British Journal for the Philosophy of Science* 24, 2, (1973), pp.153–163. However, the position adopted here is perhaps slightly stronger than that defended by Howson. Howson argued it is *possible* to assign non-zero probabilities to universal generalisations. On the view defended here, any assignment of probabilities to universal generalisations via Bayes' Theorem must be *indeterminate*. This leaves it open, however, that we might be justified in making assertions about the probability by means of some other route.

[37] It is worth here reminding ourselves of the distinction between *a priori* probability and "prior probability". The *a priori* probability of E is the probability we are entitled to ascribe to E on the basis of purely *a priori* considerations. The "prior probability" of E, on the other hand, is the probability a speaker might assign to E on the basis of general background beliefs concerning the probability of E, rather than more specific evidence bearing directly on the probability of E.

markers? Maybe the effect of the experiment will be that our apparatus, and the measuring device with it, will explode in to pieces, or melt, or be transformed in to a bunch of petunias. *Prior to us having any empirical evidence whatsoever*, these are all possibilities. Evidently, the number of such possibilities is indefinitely large. And so: the prior probability of one specific outcome, such as the needle pointing to seven, would appear to be zero. To assign to the event of the needle pointing to "7" a probability of "0.1" is not to assign to it a genuinely *prior* probability: it is an assignment of probability that assumes that there are ten and only ten possible results that might occur as a causal consequence of the experiment.[38] But *that* assumption requires background theoretical knowledge of the effects the experiment will have. And that in turn assumes knowledge of the likely effects of the experiment, knowledge of the workings of our apparatus, knowledge or belief about likely background conditions and so on.

It might be objected that even if it is granted that there are, as far as we know, infinitely many possible *effects* of any experiment, it does not follow that there are infinitely many *distinct observations* we might make after the experiment has taken place. Possibly a needle on a dial might occupy infinitely many positions after an experiment, but still, we can only discriminate between finitely many of them. Our limited powers of discrimination ensure we can only make finitely many observations.

However, it will be argued that on closer inspection this objection proves to be unsound. The objection is clearly based on the assumption that after the experiment has been performed human beings will only have finite powers of discrimination. This assumption is surely *true*, but what grounds do we now possess for saying it is true? Presumably the grounds are something like the following: we do not *now* have infinite powers of discrimination; so, probably, we will still only have finite powers of discrimination after the experiment has been performed. To say, for example, that we will not have a visual experience telling us the location of the needle to within, say, 10^{-10} cm on the grounds that we do not now possess the ability to discriminate

[38] We can perhaps get a clearer intuitive grasp of the situation by considering some "disembodied spirit" considering epistemic probabilities from behind a Rawlsian "veil of ignorance". The disembodied spirit is told that shortly it will experience *something*, it is not told what it might be. Amongst the possibilities are: the needle on a meter pointing to "7", a tropical typhoon, a cow being milked, Tony Abbott delivering a speech in Swahili and so on. There would appear to be indefinitely many possible experiences, and the disembodied spirit has, *a priori*, no reason to regard any one of them as any more likely than any other.

It might possibly be argued that there could only be finitely many possible experiences the spirit might have. An analogy might be given with a computer screen. Suppose a computer screen has N pixels and each pixel can exhibit M colours, where both N and M are finite. Then the maximum number of possible images will be N^M. This may be a very large number, but necessarily it will be finite. In a similar way, it may be argued, the spirit could only have finitely many experiences. However, it is plain that this argument is based on *assumptions* about the possible nature of the experiences the spirit will have. It is based on the assumption that the experience will be produced by some analogue of a finite number of pixels each of which can have only finitely many states. But what entitles us to make such an assumption *a priori*? From a strictly *a priori* point of view, we are not entitled to make any such assumption.

so finely, is to assess the likelihood or otherwise of this event on the basis of empirical information we in fact possess. And so it is, clearly, to assign an *a posteriori* probability to the event. And to rule out such an event by appealing to, for example, the wavelength of light or the structure of the rods and cones in our eyes is just as clearly to assign probabilities on the basis of *a posteriori* information. From a purely *a priori* point of view, we are not entitled to assert that after the experiment we will only be able to discriminate finitely many states of affairs.

It might perhaps be objected that it simply does not follow from the fact that there are, for all we know, an indefinitely or infinitely large number of *possible* results of the experiment that we must therefore assign a probability of zero to any one specific outcome. There are (at least) two ways in which this conclusion might be avoided. One way makes use of the fact that assigning a probability of zero to an event is not the same as saying it is impossible. (The probability that infinitely many fair coins, when tossed, should all come up heads is presumably zero, yet such an out-come is clearly *possible*.) So: we might assign a probability of 0.1 to the event of the needle pointing to "1", a probability of 0.1 of it pointing to "2" and so on, and assign to each one of the infinitely many other bizarre-but-possible outcomes a probability of zero. But there is a difficulty with this suggestion. On it, the probability of the needle pointing to "1" is, we are assuming 0.1, while the probability of the measuring apparatus turning in to a bowl of petunias is zero. That is, we are assuming the probability of the former event is greater than the probability of the latter. But, from a purely *a priori* point of view, it is far from clear what entitles us to assume this is so. To be sure, "common-sense" tells us it is more likely, but our common-sense is presumably here reliant on our past experience. *A priori*, we would not seem to have any reason to believe one of the events was more likely than the other. So, this way of ensuring that the probability of our observations will always be above zero, while allowing that infinitely many "outcomes" of our experiment are possible, requires us to make claims that from an *a priori* point of view are not permissible. Since we are trying to work out the *a priori* probability of "Pr(E)", this approach will not work.

There are other possible approaches, but they share the same difficulty. For example, the probability of the event of the needle pointing to "1" might be suggested to be some value less than 0.1. The same value might be assigned to the event of the needle pointing to "2", to "3", up to "10". Then the probability that either "1" or "2" or…..or "10" will be indicated will be some number less than 1; say it is N. Let 1 – N = n. Let us now consider all the other possible "outcomes": the apparatus exploding, melting, turning in to a bowl of petunias ad so on. (We will also assume that the outcomes, although infinitely many, are denumerable.) Now: order them in some way – perhaps from most plausible to least plausible. Assign to the most plausible a probability of n/2, to the second most plausible a probability of n/4, to the third n/8,…and so on. Then it will be assured that our observations will always have some greater-than-zero value, and we will have assigned to each of the infinitely many other possible outcomes some greater-than-zero value. However, the flaw in the suggestion is obvious. This suggestion requires us to say, *a priori*, that certain outcomes are more probable than others. Perhaps the outcome of the apparatus melting is assigned a higher probability than that of it turning in to a bowl of petunias. And of course, for this approach to work it is essential that *some* probabilities be greater than others. But, as already noted, from an *a priori* point of view this is not something we are entitled to say.

It might perhaps be objected that the argument just given implicitly assumes the principle of indifference, and that the status of this principle is very controversial. One statement of the principle of indifference is as follows:

If there are a number of possible outcomes of some experiment, and we have no reason to regard any one of these outcomes as either more or less likely than any other, then they ought to all be assigned the same probability. ______________________________________(PI)

The principle, *prima facie*, may seem very reasonable – even though as we noted in Chap. 2 it has its problems. It will not here be argued that PI is correct. Instead, it will be argued that the objection to Popper's argument does not rely on PI.

First, let us see why it might be thought the argument presented here relies on PI. The argument says that since we have no *a priori* reason to say that, e.g. the needle pointing to "7" is more likely (or less likely) than the apparatus turning in to a bowl of petunias, we ought to assign the same probability to both outcomes. And since the only way of doing this for the infinitely many possible outcomes is to assign each and every one of them a probability of zero, we must say the probability of the needle pointing to "7" is zero. But: this plainly makes use of the principle of indifference.

However, in order for our argument against Popper to go through, it is not necessary to use PI. It is enough to make use of the following principle, which might be called the principle of non-arbitrariness (PNA).

If there are a number of possible outcomes of an experiment and we have no reason to regard any one as either more or less likely than any other, then *we are not justified* in saying one has a higher or lower probability than any other.____________________________(PNA).

First let us note that PNA is surely unexceptionable – it comes very close to being a mere tautology. And it is also different from PI. PI tells us how we *ought* to assign probabilities. But PNA does not make any claim about how we ought to assign probabilities, it only says that certain assignments of probabilities are to be *avoided*. PI is of the form "Thou shalt", while PNA says "Thou shalt not". Still, PNA is strong enough to avoid Popper's sceptical conclusion. Any way of saying how Pr(E) could take a value greater than zero, while simultaneously allowing that, for all we know *a priori*, any experiment has infinitely many possible outcomes, must *a priori* assign a higher probability to some of those outcomes than others. And this is something that PNA says we must not do. I conclude that we do not need to use the principle of indifference in our argument against Popper.

Finally, it might perhaps be objected that this response to Popper is incompatible with the position defended in Chap 2. In that chapter it was argued that certain observations are more likely than others. How can this be if – as has just been argued – the probability of all observations is zero?

However, this objection relies on an ambiguity with the term “observation”. It has just been argued that the *a priori* probability of some unconditional observation-sentence such as “This is a black crow” being verified is zero. But our concern in the earlier chapter was with certain conditional types of probability. One claim made was that, in a universe in which all the crows are black, the probability of us blindly choosing a location for our observations in which all crows are black is one. And this claim will be true, even if the *a priori* probability of any observation being the observation of a black crow will be zero. Another claim made was that, in a universe in which some rows are black are others non-black, the probability of us blindly choosing a location in which all crows are black will be less than one. And again, this is remains true even if the *a priori* probability of any given observation being of an observation of a black crow is zero. The position adopted here with respect to the *a priori* probability of unconditional observations is compatible with the claims made in the previous chapter about probabilities that are conditional upon the universe being one way or another.[39]

---

[39] There is perhaps a more basic reason why it might be feared that in maintaining that the probability of all observations is zero we have “thrown out the baby with the bath water.” If the probability of all observations is zero, it might be thought, it must surely follow that the probability of all theories is zero too, since surely no theory can be more likely than any observation. However, this objection is not correct. On the view advocated here, the *a priori* probability of all observations is zero, and the *a priori* probability of all theories is also zero. But the *a posteriori* probability of a theory in the light of evidence need not be zero.

We can bring this out with a simple example. Suppose that a coin is tossed ten times. What is the *a priori* probability that ten heads come up? It is tempting to think that the *a priori* probability of this will be $1/2^{10}$. But, on the view advocated here, this is not so. It is, perhaps, the epistemic probability of ten heads coming up if we know in advance that the only two possible outcomes are a “heads” coming up or a “tails” coming up, and we have no reason to believe one outcome is more is more likely than another. But we do not know *a priori* that these are the only two possible outcomes. Perhaps the coin might turn in to a bowl of petunias: we do not know *a priori* that this will not happen. The *a priori* probability of a heads coming up on just one toss is, on the view advocated here, zero. So, the *a priori* probability of ten heads coming up in a row is also zero.

But now, let us assume that the coin has been tossed ten times and ten heads have come up. We now surely have good, if defeasible, reason to believe:

“The physical probability or propensity of a heads coming up on a single toss is greater than ½.”___(1)

We would have more reason to believe (1) than to believe:

“The physical probability or propensity of a heads coming up on a single toss is less than ½.”____(2)

But now, if we have more reason to believe (1) than (2), it surely follows that the *epistemic* probability of (1) must be greater than zero. Of course, (1) is not an example of what we ordinarily think of as a “theory”. But it is a non-observational statement that (together perhaps with the law of large numbers) would seem to give us probabilistic reason to believe indefinitely many other statements. The point is that, even if the *a priori* probability of all observation statements is zero, the *a posteriori* probability of a non-observational statement can be greater than zero.

Still, it may be felt, there is something “fishy” going on. On the view advocated here, the *a priori* of:

In summary, it has here been argued that we are not entitled to assert that the value of Pr(E) must be above zero. An argument was presented which seemed to indicate that Pr(E) = 0. It was noted that, as it stood, the argument was vulnerable to objection. Some ways of defending the Popperean position that Pr(E) > 0 were considered, but it was argued that all of them were flawed. More specifically, it was argued that any attempt to establish that Pr(E) > 0 violated the (apparently unexceptionable) principle PNA. It is concluded that we are not entitled to assert Pr(E) > 0, and hence that Popper's argument fails.

There is another consequence of the position adopted here that briefly deserves our attention. It has been argued that we ought to say that Pr(E) = 0. But plainly, if Pr(E) = 0, then Pr(T&E) = 0 and so, by Bayes' Theorem, Pr(T, E) = 0/0. However, the value of "0/0" is evidently "undefined" or "indeterminate". This might perhaps be thought to raise a problem for the view to be defended here. If the value of Pr(T, E) is, for any scientific theory, the indeterminate "0/0", what sense can be attached to claim that that we can have good reason to believe such a theory? And if the probability of all theories is the same, indeterminate, "0/0", what sense can be given to the claim that we can have better reason to believe one theory rather than another?

One natural response to this is to say that the considerations just given from Bayes' Theorem do not lead us to assert that the probabilities of all theories have the very same value, where that particular value has the mysterious property of being "indeterminate". It is rather that the considerations from Bayes' theorem do not assign any determinate value to any theory. This leaves open the possibility that other procedures may tell us more about the values than this route from Bayes' Theorem.

"The coin was tossed ten times and came up heads every time"_____(3)

is zero. But the epistemic probability of (1) in the light of (3), is greater than zero. But, it may be protested, this surely means that the *a priori* probability of the conditional statement:

"If the coin was tossed ten times and came up heads every time the propensity for heads to come up is greater than ½."____(4)

must be greater than zero. But, it might be thought, there is something implausible or *ad hoc* about claiming that the *a priori* probability of (3) is zero while that of (4) is greater than zero. Both (3) and (4) are synthetic, contingent statements. Why should one of them have an *a priori* probability of zero while the other does not? However, the solution to this is given by Bayes' Theorem. On the view advocated here, both (3) and "The coin has a propensity greater than ½ of coming up heads" have an *a priori* probability of zero. Hence, by Bayes' Theorem, the probability of (4) will be the indeterminate 0/0. There is, therefore, no inconsistency in maintaining that the value of (4) may be greater than zero.

3.5 The Experimentalists' Regress

Another source of doubt about the claim of science to give us good reasons for its assertions arises from the "Experimentalists' Regress".[40] In order to test a hypothesis H, a scientist must first establish that their apparatus is working correctly. But: in testing whether or not the apparatus is working correctly, the scientist may use the very hypothesis H that is supposedly under test. For example, before testing $F = ma$ a scientist might first need to test whether their apparatus for measuring force, mass and acceleration are working correctly. And in doing this, the scientist may need to make use of $F = ma$.

The experimentalist's regress might seem to suggest that – sometimes, at least – we do not really have any evidence at all for scientific hypotheses. If our criterion for determining whether the apparatus we use in testing P is working correctly is that it gives us results in accordance with P then it seems we are not subjecting P to a genuine test at all. And if P is not being subjected to a genuine test, we can hardly count it confirmed if it should pass that "test". To the extent that the theoretical claims of science are not passing genuine tests at all, we would seem to be deprived of good reasons for accepting those theoretical claims.

Like the argument from Bayes' Theorem considered in the previous section, the challenge from the Experimentalist's Regress is very general. If sound, it casts in to doubt a class of theories rather broader than those usually regarded as being of particular concern within the issue of scientific realism. As the example in the previous paragraph illustrates, it would even seem to cast in to doubt our reasons for $F = ma$. So, as with Popper's argument from Bayes' Theorem, a direct reply is needed.

It will be argued that the reasoning given in defence of the Experimentalist's Regress is faulty. Let us consider a scientist testing the hypothesis that there exists a planet in our solar system beyond the orbit of Uranus. The planet we now call Neptune was, of course, initially postulated by Adams and LeVerrier to explain apparent perturbations in the orbit of Uranus.

LeVerrier used Newton's laws of motion and law of gravitation to calculate where this planet would be. And so: the predictions made by LeVerrier of the location of the new planet were a test of, among other things, Newton's laws of motion and law of gravitation.

Suppose our scientist is aware of LeVerrier's calculations, and of his prediction that a new planet will be found in location R of space. Our scientist intends to test the prediction by looking at region R of space. But first our scientist needs to verify that his apparatus is working correctly. He does this by first observing, let us say, five other planets. He checks whether the positions of the planets, as revealed by his apparatus, is in accordance with what their locations *ought* to be at the time he observes them. But to say what their locations *ought* to be is to say what Newton's

[40] The "Experimentalist's Regress" and its significance for science, is discussed in Harry Collins *Changing Order*: *Replication and Induction in Scientific Practice*, Chicago, IL: University of Chicago Press, 1985.

laws, together with observations of "initial conditions", *predict* their location to be. If, when he uses his apparatus, the scientist finds the observed locations of the five planets to be in agreement with that predicted, he will, let us assume, conclude that his apparatus is working correctly. Clearly: in establishing in this way that his apparatus is working correctly, he has used Newton's laws of motion and gravitation. But it is Newton's laws that are to be tested by looking at region R of space to see if a new planet is there. On the face of it, this might perhaps be thought to show the supposed test of Newton's laws by looking in R forthe new planet is not a genuine test at all. But this would be a mistake.

There would be no real test of Newton's laws if the scientist's apparatus had a *disposition* to find the location of a planet to be in accord with the position predicted by Newton *no matter what the actual location of the planet was*. It is true that, when observing the five other (already known) planets, the scientist's apparatus said the planets were in the location predicted by Newton. But it does not follow that the apparatus therefore has a *disposition* to find planets to be in the location predicted by *Newton no matter what the actual location of those planets may be.* To use Newton's laws to establish the telescope is working correctly is not to establish that the telescope is *disposed* to give us observations in accordance with Newton's laws. There is at least one other possibility: the apparatus is working correctly, and it finds the planets to be in the location predicted by Newton because that is where they actually are. Moreover, the suggestion that the apparatus might find a planet to be in the location predicted by Newton *no matter what the actual location of the planet may be* is intrinsically hugely implausible. How would the apparatus – a telescope together with instrumentation for determining the direction in which it is pointing, for example – "know" what the location of the new planet predicted by Newton will be? Could the telescope somehow produce a point of light in region R even if as a matter of fact there was no planet there? Evidently, it could not. The idea that the apparatus might be disposed to give us observations in accord with Newton no matter what the actual location of the planet may be must be rejected. But: if the apparatus is not disposed to do this, pointing the telescope towards region R is a genuine test of Newton, even though Newton's laws were used in establishing that the apparatus is working correctly.

More generally, the Experimentalist's Regress need not present us with a reason for doubting that scientific theories are subjected to genuine tests. A theory T may be used to establish that some apparatus is working correctly, and then that apparatus used to test T itself. But this can still be a genuine test of T. It may be that in establishing that the apparatus is working correctly, the experimenter establishes the apparatus gives results in agreement with that predicted by T. But that does not mean that the experimenter has thereby established that the apparatus has a disposition to produce results in agreement with T *no matter what conditions actually obtain in the world.* And, we noted, the idea that the experimenter has established that such a disposition exists is massively implausible.

A defender of the skeptical argument from the Experimentalist's Regress might perhaps say that it is not the *apparatus* that has a disposition to produce only results that accord with the theory under test: it is rather the experimenter himself or herself that is disposed to do this. Perhaps the experimenter might include as "correct" observations only those that come sufficiently close to what the theory predicts, and discard the others.[41] Undoubtedly, this type of thing does go on. But the idea this is so widespread as to undermine the claim that the idea that science, typically at least, gives us good reason for its claims seems unlikely.

In summary, I conclude, the Experimentalist's Regress does not present us with a good, general reason for doubting that scientific theories are subjected to genuine tests.

3.6 The Argument from the Allegedly Unscientific Character of the Hypothesis of Scientific Realism

Gerald Doppelt has objected to scientific realism on the grounds that it fails by its own criteria of epistemic acceptability.[42] Doppelt points out that a scientific realist is prepared to accept that we have good grounds for a scientific theory only if it has predictive success, or, better, novel predictive success. But, Doppelt claims, a realist is likely to also be a naturalist, and to regard philosophical doctrines as being, broadly, the same kind of thing as scientific hypotheses. And so, a realist ought to see the doctrine of scientific realism as susceptible to the same type of empirical appraisal as scientific theories themselves. Therefore, a realist ought to see scientific realism as an acceptable hypothesis only if it itself has had predictive success, or, ideally, novel predictive success. But, claims Doppelt, scientific realism has not had success of that type, and so by the realist's own epistemic standards ought not to be accepted.

It is useful here to note that "scientific realist theses" can be more or less specific. A *general* scientific realist thesis might be "Most of contemporary science is (roughly) true and the entities postulated by it exist." A more *specific* realist thesis might be "Atoms exist and behave more or less as our theories say they do." We might, perhaps, discover that some of the entities postulated by current theory do not exist. For example, we might discover that "strings" do not exist. But: very many more specific realist theses might remain unrefuted.

Now, let us focus on one example of a more specific realist thesis:

Atoms exist and behave more or less as our theory says. _____(1)

---

[41] Perhaps one of the most famous alleged examples of a scientist using only the data that seemed to favour their preferred theory was Eddington's observations in 1919 of an eclipse as observed from the island of Principe. However, recent investigations have found that Eddington did not selectively use data. See Daniel Kennefick "Not Only Because of Theory: Dyson, Eddington and the Competing Myths of the 1919 Eclipse Expedition" Cornell University Library (2007), http://arxiv.org/abs/0709.0685

[42] See Gerald Doppelt "Empirical Success or Explanatory Success: What Does Current Scientific Realism Need to Explain?" *Philosophy of Science*, v.72, (2005), pp.1076–1087.

If (1) is to be taken as a realist thesis, then, surely, it is to be understood as literally true, and not given some non-realist epistemic or operationalist interpretation such as:

“Atoms exist and behave more or less as our theory says” has passed a wide variety of tests.___________________________(2)

But, if (1) is given a realist interpretation, then it surely *does* lead to a class of novel predictions. It leads to the predictions that our theory of atoms will successfully pass any *new* tests we throw at it. Moreover, (2) does not lead to any such novel predictions. From the fact that our theory of atoms has passed tests so far, it does not logically follow that it will pass new tests to which we subject it. The specific realist thesis (1) leads to a novel prediction, whereas its non-realist counterpart (2) does not.

Note that there is nothing in the above argument that requires us to assume that truth is a causal-explanatory concept; neither does it require us to accept any philosophically tendentious metaphysical-realist concept of truth. It only requires us to assume that when a realist asserts (1), they are asserting something more than (2). In particular, it assumes that to adopt a realist view of (1) is to say (1) will pass new types of tests, not just the ones it has passed so far.

It might be protested that there are also some non-realist interpretations of (1) that also lead to novel predictions:

“Atoms exist and behave more or less as our theory says.” has passed all our tests, and will pass the next type of test to which we subject it.____(3)

(3) makes a novel prediction, but surely falls short of expressing a fully realistic view of atoms. However, this does not show realism does not make novel predictions. It merely shows that it is *not the only* philosophical interpretation of science that is capable of doing so. We return to the significance of this shortly.

There are of course, some very strong realist theses that make novel predictions that have been falsified. Consider, for example:

All scientific theories that have ever been advanced are true.__(4)

This makes the prediction, or retrodiction, that any scientific theory that has ever been advanced would pass any test thrown at it. And this very strong thesis has of course been shown to be false. But the fact that such a thesis has been *falsified* would appear to confirm its scientific status. Scientific realist theses are *falsifiable*, and some of them have been falsified. But that does not mean all of them have been falsified. For example, (1) has not been falsified; neither, surely, has the claim that most entities of mature sciences exist and behave more or less as our theories say.

It is perhaps natural to protest that in the example given, all the explaining is done by the hypothesis of atoms itself. The philosophical theory of realism is otiose: a scientist could explain the ability of the theory of atoms to pass the new test in purely scientific, empirical terms, without reference to scientific realism.

However, it will be argued that this objection is not quite right. First, let us remind ourselves of a point already made: if our theory of atoms is given the non-realist, operationalist-style interpretation (2), then it ceases to make the prediction it will pass *new* types of test. But if it is given a realist interpretation, it does make this prediction. So there seems to be a sense in which adopting realism does make it possible to derive predictions that otherwise could not (or might not) be possible to derive. So, there does seem to be a sense in which realism *contributes* to the derivation of the prediction. But, the way in which it contributes to the making of the prediction is *not by being conjoined* with the scientific, empirical theory of atoms. (Its contribution is not like that of an auxiliary background hypothesis that is *conjoined* with the theory to make possible the derivation of the prediction.) Scientific realism makes possible the derivation of the prediction by instructing us to *interpret* the scientific theory in a particular way. If the theory is interpreted as the scientific realist recommends, the prediction can be made, but if it is interpreted in (some) non-realist ways, the prediction cannot be made.

It might perhaps be protested that we do not need a *realist* interpretation to be able to get out the prediction that the theory will pass a new test. All we need to do to be able to derive the prediction is simply *refrain* from adopting some non-realist, epistemic interpretation of the content of the theory, and just take the theory at "face-value". And this may be correct. A scientific realist (as opposed to a *metaphysical* realist, for example) may be quite content to simply say that the claims, including the existence claims, of scientific theories about unobservable entities are to be taken at face-value.[43]

Let us return to a point noted earlier: scientific realist thesis (2) makes a novel prediction, but so does the non-realist thesis (3). What reason might we then have for preferring realism to the non-realist interpretation expressed in (3)? A tempting first response might be that (2) provides a *better explanation* of novel success than (3), and on those grounds it is more worthy of rational acceptance. However, it is clear that such a response assumes the soundness of inference to the best explanation. In Chap. 4 it is argued we are not entitled to make that assumption. So, we are confronted with the question: Do we have any reason to assert the bolder (2), rather than merely assert the logically more modest (3)?

One of the main themes of this book is that we do have such good reasons. Eddington inferences, it will be argued, entitle us to say we have probabilistic reasons for accepting realist claims such as (2).

[43] A similar point is made by Alan Musgrave is response to Arthur Fine's defence of the "natural ontological attitude". See Musgrave "The Ultimate Argument for Scientific Realism" in Robert Nola (ed), *Relativism and Realism in Science* (Kluwer Academic Publishers, 1988), pp.229–252.

3.7 The Theory Laden-Ness of Observation

On the view to be defended here, we do not arrive at realist claims by inferring them (deductively or otherwise) from our best theory, but by means of Eddington inferences from observations sentences. Eddington inferences, it will be argued, can be given a probabilistic justification. So, on the view to be defended here, we arrive at realist claims by probabilistic inference *from observation sentences*. On this approach, it does not matter whether or not we have good reason to believe our current *best* theory. But there is another point at which the route to realism to be used here *might* be vulnerable to possible objection. We arrive at realist claims by inference from observation sentences. The inference is of the form: "If observation sentences O are true, then, probably, realist claim R is true". But, plainly, this will give us good reason for realist claim R only if we have good reason to believe the observation sentences O. And this is what the theory laden-ness of observation might seem to cast in to doubt.

The expression "the theory laden-ness of observation" refers to a cluster of related phenomena which, in one way or another, seem to show that observation-statements do not constitute an absolutely firm foundation for the construction of scientific theories. Sometimes what we expect to see influences what we report seeming to see. Plausibly, an example of something like this phenomenon is common in everyday life. We expect to see the cat on the floor. We see a dark shape there out of the corner of our eye. In fact, it is a sock, but somehow we seem to experience the visual sensation of a cat. Or again, the way we classify an object can be determined by our "way of seeing", or by the "gestalt" we bring to our perception.[44] The animal looks like a rabbit, but perhaps we are merely seeing it as a rabbit and it is actually a duck. Sometimes our prior beliefs or systems of classification or commitment to a scientific paradigm influence how we perceive things.

Perhaps the first thing to note here is that while, of course, there can be reasons for doubting ordinary observation statements, it is not entirely clear why this ought to be seen as undermining the case for realism. Realism is, I think, it is fair to say, generally opposed to Instrumentalism, or, perhaps to "Constructive Empiricism". Opponents of realism say that we are not epistemically entitled to say that claims about unobservables are true, or that statements about unobservables are not the type of thing that are capable of being true. An instrumentalist, for example, will say we ought to restrict our assertions to things like needles on dials pointing to numbers on meters, and the like. But instrumentalism is different from an extreme, blanket scepticism. An instrumentalist will allow that we can have good reason to believe

[44] The notion of the theory laden-ness of observation was perhaps raised in prominence as a topic of discussion in the philosophy of science by Thomas Kuhn's *The Structure of Scientific Revolutions*, esp. p.111, pp.113–114, pp.115–121. It is not claimed here that the considerations given by Kuhn fail to show that observation statements do have a risky and theory-laden character. Rather, the claim is that even though they do have that character, Moorean arguments show they are firmer than theoretical claims, and that under ordinary conditions of utterance they are things we can be said to know.

that the meter, for example, exists: they will (typically) not say that only the perception of the meter exists, or that it is merely an impression created in our minds by God. Realists and instrumentalists typically agree that we can have good reason for assertions about ordinary-sized material objects – what divides them is the question of whether we can also have good reason to believe in the existence of unobservable entities. So: while *there are* reasons for doubting statements about familiar observables, such reasons for doubt would not seem, in themselves, to be grounds for favouring instrumentalism over realism since they also apply to instrumentalism. They do not count for instrumentalism *over* realism. Neither would they seem to count in favour of other rivals to realism such as constructive empiricism.

3.8 The Objection from Unconceived Possibilities

P. Kyle Stanford has produced an extended defence of a version of the argument against realism from unconceived possibilities.[45] Briefly, this argument is as follows. At any one point in time, scientists select the best available theory. But the best available theory is only selected from the "pool" of theories that scientists have, at that point in time, been able to come up with. And so, the possibility arises: perhaps there might be some better, as yet unconceived-of theory that offers a superior alternative explanation of the data. This consideration leads to an argument against the idea that we can have good reasons for realist claims.

Let T be the best theory, explaining known data, which scientists have been able to come up with. But suppose there is some better, as yet unconceived theory T*, that can explain that same data. If, hypothetically, T* were known to be better than T, it would hardly seem rational to adopt realism with respect to T. And so, the fact there *may*, for all we know, be some better but as yet unconceived-of theory would seem to at least to some extent undermine the case for realism about T.

However, it might be protested that the sheer fact that *for all we know* there might be some unconceived-of better theory hardly seriously undermines the case for realism about T. Perhaps the chances of there being some unconceived-of better theory are very low. Stanford responds to this type of criticism by arguing that the history of science gives credence to the idea there probably are better but as yet unconceived of theories. He argues that this presents a more serious challenge to realism than either the pessimistic meta-induction or the argument from the underdetermination of theory by data.

Clearly, Stanford's argument assumes realism is the thesis that we ought to adopt a realist view of the *best* theory. But in this book that thesis is rejected. Instead, it is argued we are entitled to adopt a realist view *of certain claims only if we are led to them by Eddington-inferences*. If we are not led to the entities postulated by our best

[45] See P. Kyle Stanford *Exceeding Our Grasp: Science, History and the Problem of Unconceived Alternatives* (Oxford University Press, 2006)

theory by an Eddington-inference, then no claim is made here that we ought to be realist about those entities.

It is worth examining in a little more detail why Stanford's argument would seem to have much less force against the view advocated here than it might against IBE-related arguments for realism. Let us begin by reviewing the nature of Eddington inferences. We make an Eddington inference when we go, for example, from:

(1) The fish in our nets are all 2 in. or larger, and the size of the holes in our fish net are exactly 2 in. across.

To the conclusion:

(2) There are (probably) fish in the sea of sizes smaller than 2 in.

When (2) tells us that there probably smaller *fish*, it is, we may assume, making the claim that there are probably in the sea creatures *resembling* those in our nets – that is, creatures possessing gills, fins, of a certain body shape and so on – but *smaller* than those in our nets.

The statement (2) does not rule out the possibility of there being in the sea other "weird and wonderful" creatures quite different from those we find in our nets: it merely says that, probably, there exist creatures similar to those in our nets, but smaller. So, the above Eddington inference does not *rule out*:

(3) There exist in the sea unconceived-of creatures.

Although (3) is not ruled out, it seems intuitively clear that (2) is directly supported by (1), in a way that (3) is not. The reason why (1) directly supports (2) is, roughly, as follows: Suppose the smallest fish in the sea was exactly 2 in. long. Then a highly improbable event would have occurred: the size of the smallest fish in the sea would have happened to have coincided with the size of the holes in our fish net. Since such a thing seems highly unlikely, we may conclude it has probably not occurred, and therefore that the size of the smallest fish in the sea is not 2 in. That is, there are probably fish – that is, creatures *similar* to those in our nets – of a size smaller than 2 in.

This line of reasoning supports the existence of things smaller than but otherwise similar to those in our nets. But it provides no support at all for the existence of things *unlike* fish. It provides no support for the existence of "unconceived-of" creatures. They *may* exist, but their existence is not supported by the Eddington inference.

The aim of the later chapters of this book is to argue that some, but not all, realist claims can be supported by Eddington-inferences. If a type of theoretical entity is supported by an Eddington inference then, it is argued, we have probabilistic reason to believe entities of that sort exist. But, it will also be argued, the sheer fact that a theory is the best theory is *not* sufficient for us to be realists about it. Eddington-inferences, it will be argued, are necessary for rational belief.

If the point of view to be defended here is correct, then the possible or even likely existence of unconceived-of but better theories than those already known poses no threat to realism. Even if such possible but unknown theories were better than those that scientists have managed to come up with, it does not follow that the case for realism with respect to such theories would thereby be stronger. Such theories would also need to have their existence claims supported by Eddington inferences. Moreover, on the view to be defended here, the existence of Eddington inferences to such entities (without either or a good or bad theory postulating them) is sufficient to justify belief in the likely existence those entities.

These issues are discussed in greater length in Chap. 5. But in this section reasons have been sketched why the argument from the possibility of unconceived-of theories would not appear to endanger the type of realism advocated here.

3.9 Concluding Remarks

The aim of this chapter has been to defend the idea that science gives us good reason for its theoretical claims against some influential attacks. We can briefly recapitulate some of the main points.

It was argued that the Pessimistic Meta-Induction at most only threatens one aspect or dimension of realism. The thesis of Scientific Realism, we noted, had two dimensions: the behaviour dimension and the existence dimension. The Pessimistic Meta-Induction would not seem to endanger the behaviour dimension. It might, however, seem to place the existence dimension under some threat. But: we noted that defending the existence dimension of Scientific Realism is precisely what is done by Eddington inferences.

The underdetermination of theory by data has been another source of scepticism about realism. But the underdetermination thesis only casts in to doubt the more theoretical parts of science if we make additional assumption that empirical evidence is the *only* form of evidence for a theory. One of the aims of this book is to argue the assumption is not true: it is argued here that empirical evidence is *not* the only form of evidence. And so it is a consequence of the position adopted here that underdetermination need not threaten realism.

A number of philosophers, including Laudan, Leplin and Stanford have also argued that underdetermination does not threaten realism. These authors approach the matter from a very different angle from the one adopted here: none of them appeal to the possibility of non-empirical evidence or to what we are here calling Eddington inferences to sever the link between underdetermination and anti-realism. But it was argued that the approaches of these other authors do not succeed. We still need Eddington inferences to establish the case for Scientific Realism.

The argument from Bayes' Theorem purporting to show that the probability of our theories always remains at zero was discussed. It was argued that the considerations from Bayes' theorem actually do not permit us to assign any determinate value to the probability of our theories.

It was argued that it is a mistake to think considerations from the "Experimentalist's Regress" show theories are not confirmed. It was argued that even if a theory is used in establishing that apparatus is working correctly, still the theory may be subject to genuine test and confirmation.

The argument from the theory-dependence of observation failed to give us good reason to reject realism. For the purpose of defending realism, it is sufficient to show that the case for some unobservable entities is as good as that for observable entities.

Finally, it was argued that the objection from unconceived-of possibilities need pose no threat to the view adopted here.

The last two chapters have examined a range of arguments against the idea that we can have good reasons for scientific realism. Inductive scepticism was considered in Chap. 2. In this chapter a range of other arguments have been considered. It has been argued that none of these arguments threaten to damage the specific way of arguing for realism to be defended in this book.

Chapter 4
Realism and Inference to the Best Explanation

The overall aim of this book is to develop an account of the nature of the reasons for scientific realism. So far, our aims have been negative or defensive: they have been to show that the (sceptical) arguments *against* the rationality of realism are (mostly) flawed. In this chapter we begin task of exploring the arguments that have been advanced *for* realism.

Perhaps the most widely used of the arguments for Scientific Realism is that it provides the *best explanation* for the success of science. Forms of this argument have been defended by J. J. C. Smart, Hilary Putnam, Richard Boyd, Alan Musgrave, Peter Lipton, Jarrett Leplin, Stathis Psillos and others.[1] The aim of this chapter is to critically examine this type of argument as a route to Scientific Realism. Briefly, it will be argued that whether the argument is good or not depends on how the expression "best" is understood. Again, very briefly, if "the best explanation" is taken to be "the simplest explanation" or something similar to "simplest", then the argument is not a good one. There are other possible interpretations of "best" on which the argument may be good, but it is argued it is unclear whether these interpretations of "best" are enough to *justify* a belief in scientific realism, as that doctrine is normally understood.

[1] See J. J. C. Smart, *Philosophy and Scientific Realism* (Routledge, 1963), Hilary Putnam, *Matter, Mathematics and Method*, p. 73) Richard Boyd "On the Current Status of the Issue of Scientific Realism" in *Erkenntnis*, 19, pp. 45–90., Alan Musgrave "The Ultimate Argument for Scientific Realism" in R. Nola (ed), *Relativism and Realism in Sciences*, (Dordrecht, Kluwer), Jarrett Leplin *A Novel Argument for Scientific Realism*, (OUP, 1997), Stathis Psillos, *Scientific Realism: How Science Tracks the Truth* (London, Routledge, 1999.)

J. Wright, *An Epistemic Foundation for Scientific Realism*, Synthese Library 402, https://doi.org/10.1007/978-3-030-02218-1_4

4.1 Some Preliminary Issues

Scientific Realism will here be understood as the thesis that the entities, including the theoretical or unobservable entities, postulated by scientific theories, exist and behave (at least more or less) as those theories say they do. Of course, this simple statement of Scientific Realism might be qualified in various ways. Some Realists might not wish to adopt a realist stance with respect to *all* scientific theories: they might, for example, only wish to be realist about "mature" scientific theories. Some might only wish to be realists about those theories that have had *novel* predictive success. Some might only wish to be realists about the "structure" described by the theories while refraining from making any claim about the "nature" of the entities dealt with by the theories. Still others might only wish to be realist about those parts of a theory somehow "directly deployed" in making (novel) predictions. These various refinements to or modifications of Scientific Realism will be considered in due course. For the moment, however, we will adopt as our working definition the one given at the head of this section: the entities postulated exist and behave more or less as the theory says they do.

Arguments for Scientific Realism from Inference to the Best Explanation (IBE) can be divided in to two types. These can be represented as follows:

Type One: Arguments from the fact that since specific scientific theories provide the best explanation of the phenomena with which they are concerned, we ought to be realists with respect to the subject matter of those theories. (For example, if a given theory of atoms gives us the best explanation of certain observable phenomena, we therefore ought to be realists about atoms.)

Type Two: The argument from the fact that since the hypothesis of Scientific Realism itself gives us the best explanation of the success of science, we ought to accept as true the hypothesis of Scientific Realism.

In the arguments of the first type, the explananda are various empirical phenomena that have been observed by scientists and each explanans is a specific scientific theory, realistically interpreted. But in the second type of argument the explanandum is the success of science in *predicting* those empirical phenomena, and the explanans is the philosophical theory of Scientific Realism itself. However, both arguments use an assumption: that if a theory provides the best explanation of some fact or facts, we are thereby rationally justified in asserting that explanation to be true. Is this an assumption we are entitled to make?

4.2 The Accessibility of the Fact That a Theory Is "the Best"

Clearly, whether or not we are entitled to make the inference from "This theory is the best" to "This theory is (probably) true" will depend on what, exactly, we take "best" to mean. But, however "best" is to be interpreted, there is one constraint that must be placed on any interpretation of this term. The property of being "best" – however that term may be construed more precisely – must be taken to be an *accessible* property of theories.

One way we can bring this out is as follows. Suppose we were to say that a theory is "the best" only if it is, in fact, true. Then, we would certainly be (trivially) rationally entitled to make an inference from "T is the best explanation" to "T is true". But such an interpretation of "best" would seem to defeat the whole purpose of "inference to the best explanation". Whether or not a theory is true is – at least if it makes claims about unobservable entities – is not something we can determine by direct observation. We may not, for example, be able to directly peer in to less accessible parts of reality, such as the interior of the atom or remote reaches of space and time. So, there would seem to be a sense in which we are unable to "tell directly" whether or not theories in these domains are true. Compared to observable states of affairs, the truth or otherwise about these domains is *relatively inaccessible*. But, if the theory about how things are in these less accessible domains is *the best* available, then, we hope, this fact gives us good reason for belief that those theories are (probably) true. But then, of course, if a theory's being *the best* is to play that role, the fact that it is the best must be *more accessible* than the fact that it is true.

There is another line of thought that brings out a how a theory's being "the best" must be an accessible fact about that theory. The idea of a theory being "best" is related to it being good: more specifically, to say a theory is best is to say it has *more goodness than the others*. And to, to say it is good, or has more goodness, implies that there is a sense in which we *ought* to prefer it because it is the best. "Ought" implies "can". Hence, to say a theory is the best seems to imply that there is a sense in which we are *capable* of preferring the best theory to the others. But this in turn implies that we are capable of *recognising* a theory as the best when we encounter it, and hence that the fact that a theory *is* the best needs to be something that is accessible to us.

4.3 Probability

It is worth beginning with what might seem to some to be the most natural candidate for the property we want: probability. Might the "best theory" be understood as the "most probable theory"?

It is customary to distinguish between epistemic probability and objective probability or propensity. To say there is an objective physical propensity for something to be the case is to make a claim about how things are in the empirical world. It is a claim the truth-value of which has the same type of inaccessibility as the claims of empirical science themselves. So, to say that a theory is the best because it has a sufficiently high propensity to be true would fail to meet the requirement of accessibility defended in the previous section.

A related point can be made if statements about probability are interpreted as statements about frequencies. Consider the assertion:

In all possible states of the world in which observation statements O hold true, theory T is true n% of the time.________________(1)

Even if there are statements of the form of (1) that are in fact true, the sheer *fact* of their truth is not enough to furnish us with good *reasons* to believe T. We need *reasons* for claims such as (1). And, at least if T makes claims about unobservables, we have yet to see what those reasons might be.

Perhaps then *epistemic* probability is the notion we need. Perhaps a theory T is to be counted as the best if it has a higher epistemic probability than its rivals.

Unfortunately, things are not quite that simple. If we have no reason to regard a "Heads" as more likely than a "Tails" when a coin is tossed, then plausibly "Heads" and "Tails" have the same epistemic probability. In this case, perhaps, the *absence* of a reason makes it the case that two outcomes are equally likely.[2] But if we are to assert that a "Heads" for example, has a *higher* epistemic probability than a "Tails", then we must have *some positive reason* for assigning a higher probability to "Heads". We must be able to give some answer to the question: "*Why* does "Heads" have a higher epistemic probability than "Tails"?" And the answer will be of the form: "Heads has a higher probability than tails *because of* ABC", where "ABC" refers to something other than the mere fact that heads has a higher epistemic probability than tails. Also, for reasons already given, "ABC" needs to be *accessible*.

Although epistemic probability would certainly seem to typically be more accessible than propensity, still, any claim that one theory has a higher epistemic probability than another needs to be supported by a *reason*. And so, we are led to the question: what type of thing can constitute a reason for saying one theory has a higher epistemic probability than another?

4.4 Simplicity

One natural candidate for a property that confers a higher epistemic probability on a theory is simplicity. But: are we actually entitled to make an inference "T_1 is simpler than T_2" to "T_1 is more likely to be true than T_2"? Do we possess a *justification of simplicity*?

In considering this question, it is worth getting clear on precisely what is at issue. We are considering inference to the best explanation as a route to Scientific Realism. Therefore, we are considering whether the simplicity of a theory is a good reason to say it is true in the "full-blooded" realist sense of making true claims about the world at *both* the observational and theoretical level. To say a theory is "true" in this sense is not merely to say that it can explain the observed phenomena, or even that all of the observation sentences it (perhaps in conjunction with other true statements) implies are true. It is to say that both its claims about what we will observe, *and its claims about what is going on at the non-observational*, *theoretical level*, are

[2] Again, we here overlook the difficulties associated with the principle of indifference.

true. It will be argued that we do not seem to be in possession of a good argument for the claim that if a theory is simple then it is true or likely to be true in this "full-blooded" realist sense. And clearly: if we do not have such a "full-blooded" justification of simplicity, then any argument for realism that appeals to IBE would seem to be undermined.

4.5 Simplicity and Curve-Fitting

Considerable work has been done on applying the notion of simplicity to "curve-fitting". An indefinitely large number of curves can be drawn through, and extended beyond, any finite number of data points on a graph. In practice, we tend to prefer the smoothest or simplest curve that passes tolerably close to the data points. But although we in practice tend to prefer the smoothest, simplest curve, and intuitively regard it as more likely to be correct, the question arises: Do we actually have good reason to prefer the simplest curve?

One recently influential approach to these issues argues we have reason to prefer curves the equations for which minimise the number of freely adjustable parameters.[3] Briefly, the argument for the preferability of such curves is as follows. Suppose we have some collection D of data points. If we are permitted to use equations with indefinitely many freely adjustable parameters, it would hardly be surprising if we managed to find *some* curve that passed through (or tolerably close to) all of the data points in D. We would, therefore, have no grounds for saying that such a curve would tend to yield correct *predictions* about what observations we would get beyond the actually obtained data D. But finding a curve with a small number of freely adjustable parameters that nonetheless managed to pass through or very close to all the data points would be a real achievement. There would in such a case be more reason to think we really had got on to the "correct" curve; where to say the curve is "correct" is at least to say it would correctly predict new observations. Two currently influential criteria for selecting the simplest curve are the Akaike Information Criterion (AIC)[4] and the Bayesian Information Criterion (BIC).[5]

---

[3] One influential account of this sort is developed in H. Jeffreys *Theory of Probability* (Oxford, Clarendon, 1961). His initial account applied to laws expressed as differential equations. He argued that, given that some candidate laws expressed as differential equations can all account for some body of data D, the more complex the differential equation, the less likely it is that it should be the correct law. The complexity of the differential equation E is given by $K = a + b + c$, where a = the order of the differential equation, b = its degree and c = the absolute value of the coefficients in it. Then, Jeffreys argued, the probability that E is correct is given by 2^{-K}.

[4] See H. Akaike "Information Theory and an Extension of the Maximum Likelihood Principle" in Petrov, B. N. and Csaki, F. *2nd International Symposium on Information Theory*, Armenia, 1971, Akademiai Kiado, pp. 267–281.

[5] See Gideon E. Schwarz "Estimating the dimension of a model" in *Annals of Statistics* **6** (2), pp. 451–464. (1978).

However, despite the great interest of such developments, it is clear they are of no help in the present context. That they are of no help can be brought out by the following dilemma. The data points in D, we may suppose, tell us the magnitudes of sets of properties. For definiteness, let us assume our data points are a number of pairs $<p_a, p_b>$ that tell us the magnitudes, at various locations, of two properties P_A and $P_{B.}$ The dilemma arises when we consider the question: "Are the properties P_A and P_B observational properties or theoretical properties? Suppose they are observational properties. Then the techniques for choosing the simplest curve that passes through the data-points will only tell us about the likelihood of various claims about other, so far unobserved values of these observable properties. It will not tell us anything at all about the likelihood or otherwise of theories about how unobservables might be causing those values. For example, suppose we are observing the behaviour of a gas, and P_A is temperature while P_B is pressure. Then criteria such AIK and BIC will tell us certain values for temperature and pressure are more likely to be obtained for new data. But these criteria will tell us nothing about the likelihood or otherwise of the suggestion that the observed relations between temperature and pressure are due to the motions of molecules.

Now, let us suppose that P_A and P_B are non-observational or theoretical properties. Then techniques such as AIC and BIC of course will lead us to say that, *given the obtained values for the properties P_A and P_B*, some theories about the behaviour of those unobservables are more likely to be true than others. But now, suppose both criteria tell us that, *given the assumption that we have already obtained some true values of the unobservable properties P_A and P_B*, theory T_1 is more likely than theory T_2 to yield true claims about new values for those properties. Would this mean we are in a position to assert that T_1 is more likely to be true than T_B? Plainly, we are not yet in a position to make such a claim. P_A and P_B, we are assuming, are *non*-observational properties. And so the questions arise: "What grounds do we have for assuming that the values of P_A and P_B used to make the prediction are *correct*?" "What grounds do we have for saying there even exist any such non-observational, theoretical properties as T_A and T_B?" The techniques AIC and BIC are of course no help in answering these questions: they only enable us to extrapolate beyond initial to values to new values; they do not tell us whether the initial values are correct or even constitute measurements of a real property. The problem of determining whether or not we have good reason to believe assertions about unobservable entities remains untouched.[6]

Nonetheless, the progress that has been made in providing a probabilistic justification of curve-fitting practice might give us some degree of optimism about developing a probabilistic route to realism. After all, it (curve-fitting) is a type of ampliative inference that relies on something like simplicity rather than mere

[6] The argument given assumes that either both properties are observational or both are theoretical. It is, of course, possible, that one is observational and the other is theoretical. However, it is plain that this would not help us to establish realist claims. If one of them (P_b, say) is theoretical, then the question would arise: how can we establish its values. And: if we cannot, we can use criteria such as AIC to establish the mostly likely relationship between the variables.

enumerative induction. If it can be given a probabilistic justification, maybe other types of inference that get us closer to realism can also be given such a justification. We return to these themes in later chapters.

4.6 Could Appeal to Simplicity Justify Realism?: Some General Remarks

In the previous section we noted that there were curve fitting techniques that gave us good grounds for thinking that simpler curves were more likely to be true than less simple ones. So, within this limited context, there can exist a probabilistic justification of simplicity. However, it was also noted that this limited justification of simplicity was no help in getting us to realism. In this section we consider just what would need to be done if we were to justify realism via the notion of simplicity. It is useful to distinguish the following two questions:

(i) Might a theory T that postulates unobservables be simpler than rival theories explaining the same phenomena that do not postulate unobservables?
(ii) Might the fact that T is simpler than its rivals also mean that T is more worthy of rational acceptance than its rivals?

It does seem to be at least *possible* for a theory that postulates unobservables to be simpler than another theory, explaining the same phenomena, that does not. Consider the theory that matter is made of molecules that obey Newton's laws of motion. This provides us with a unified explanation of an array of disparate phenomena including the gas laws, mixing and diffusion phenomena and Brownian motion. What would otherwise be a number of distinct laws and generalisations are seen as differing manifestations of the same underlying physical processes. We might, perhaps, appear to endure some loss in simplicity in postulating molecules, but, it could with some plausibility be maintained, we actually get an overall gain in simplicity as a result of the unified explanation we are able to construct. And so, this example seems to show it *may* be possible for a theory that postulates unobservables to be simpler than one that does not.

But, of course, the above considerations do not by themselves quite furnish us with a good argument for realism. In order to have such an argument, we need to address (ii), above: Is there reason to say that the greater simplicity of a theory *increases its degree of rational credibility*? We have yet to see how that might be so.

There is one very clear type of case in which a theory T can be more likely, and hence more rationally credible, than another theory T*. Suppose T is equivalent to "p" while T* is equivalent to "p&q", where "p" does not entail "q". Then T will intuitively be simpler than T*. But it is also very clear that T must be more likely than T*. Might this familiar feature of probability help us to develop a case for realism? It will be argued that there are two obstacles to such an approach.

First, it seems very clear that it is not possible for a theory postulating unobservables to be of the form "p" while a theory postulating only observables is of the form "p&q". If our theory T_O postulating unobservables is of the form "p&q", then any theory of the form "p" will, necessarily, also only postulate unobservables. If our theory T_O, of the form "p&q", is a conjunction of laws that relate observable properties to observable properties, then any theory of the form "p" will just be a smaller set of those laws relating observable properties to observable properties. It cannot be about unobservable, theoretical entities. Therefore, this approach does not provide a mechanism whereby a theory postulating unobservables might be more likely than one postulating only observables.

Still, it might be thought this general approach has promise. Let us remind ourselves of just how "p&q" is less likely than "p". Assuming "p" and "q" are independent, Pr(p&q) is given by Pr(p) x Pr(q). And since probabilities always take a value no greater than one, Pr(p) x Pr(q) cannot be any greater than Pr(p). Unless Pr(q) = 1 – which presumably will never be the case if "q" is an empirical generalisation or explanatory law – Pr(p) x Pr(q) will be less than Pr(p). Now, these considerations suggest a possible way in which it might be shown that a theory postulating unobservables could be more likely than one postulating only observables. Let us suppose, purely hypothetically, that we had good reason to believe all assertions of a certain class had the same likelihood of being correct, irrespective of their specific content. (For example, suppose hypothetically that we had good reason to believe all assertions of the form "All As are Bs" had the same likelihood of being true, irrespective of the meaning of "A" and "B".) We will call such assertions "basic assertions". Suppose a theory containing only observable terms could be expressed as the conjunction of N such basic assertions, while a theory containing terms for unobservables could be expressed as the conjunction of M basic assertions, where $M < N$. Under such circumstances, we would, surely, have good grounds for saying the theory postulating unobservables was more likely.

Unfortunately, however, this suggestion is not really of much help. It is based, clearly, on the assumption we can identify some class of statements each member of which has the same likelihood of being true, and that statements of this sort are sufficient to express (typical, at least) scientific theories. Obviously, showing that is the case would be a large undertaking. But there is one particularly difficult obstacle that such an approach would need to overcome. As we are interpreting it here, scientific realism has both an existence dimension and a behaviour dimension. It will be argued that a serious problem confronting the approach currently under consideration arises from the existence dimension of scientific realism.

Suppose we are faced with the choice between: (i) adopting a realist view of atoms and molecules as an explanation of the gas laws, and (ii) merely believing the gas laws themselves, considered as laws relating the quantities pressure, volume and temperature. Since realism, on the view adopted here has an existence dimension, (i) will entail:

Atoms and molecules exist.________________________(2)

But (ii) on the other hand, will only commit us to:

Gases exist.__________________________________(3)

Although (2) and (3) are both existential statements, their epistemic probability *in the light of the observations we have made* will be very different. The likelihood of (3) will presumably be very high. But what will be the likelihood of (2)? On the face of it, its epistemic probability would surely appear to be rather lower than that of (3). But, the approach currently under consideration would not seem to even suggest a way of determining the likelihood of (2). And this *obviously* creates a difficulty. The approach under consideration requires us to be able to identify classes of statements with the same probability, and choose the conjunction with the fewest members capable of explaining the data. But, if we are to have good reason to accept realism, this will require us to choose between competing existential claims, where some of these existential claims are about observables, and others about unobservables. Such competing existential claims will pretty evidently *not* have the same probability, and we have as yet no idea how to say how probable the claims about unobservables might be.

Clearly, if the approach being considered is to work, we would need to already be in possession of some method of assessing the likelihood of existential claims about unobservables. But, if that is so, we are back at square one: our overall aim is *develop* some way of determining the rational credibility of existential claims about unobservable entities.

In summary, there is one obvious sense in which in which simple theories are more probable than theories that are not simple: "p" is simpler than "p&q", and is also more probable. But it has been argued that this seems unlikely to be able to furnish us with a justification of *realism*. It seems to require of us precisely what we are after. What is perhaps the most obvious way of justifying simplicity does not seem to work.

Many authors have given a wide range of reasons for, in one sense or other of the terms, *preferring* or *using* simple theories. For example, Karl Popper has argued that simple theories tend to be more falsifiable, and high falsifiability is to preferred, on his view, since it hastens the progress of science.[7] Michael Friedman has argued that increasing simplicity is desirable because it increases our *understanding*.[8] Will Derske has defended the use of simplicity on aesthetic grounds.[9] David Lewis has defended something like the thesis that perhaps it is a basic or fundamental fact that simplicity is a property of theories that it is appropriate to value.[10] But none of these

[7] See K. Popper *The Logic of Scientific Discovery*, Section 43.

[8] See Michael Friedman "Explanation and Scientific Understanding" in *The Journal of Philosophy* **71** (1974), pp. 5–19.

[9] See W. Derske *On Simplicity and Elegance* (Eburon Press, 1992). Similar ideas, relating aesthetic value to simplification or unity of complexity (together with intensity), were developed by Monroe Beardsley in *Aesthetics*: *Problems in the Philosophy of Criticism* (Harcourt, Brace and World, 1958).

[10] See David Lewis *Counterfactuals* (Basil Blackwell, 1973), p. 87 A similar sentiment was once expressed by Elliot Sober "What is the Problem of Simplicity?" in Zellner, A., Keuzenkamp, H. McAleer, M. *Simplicity*, *Inference and Modelling* (Cambridge University Press, 2001), pp. 13–31.

positions constitute what we would ordinarily be inclined to count as a *justification* of simplicity, in the sense of an argument, starting from widely accepted or reasonably uncontroversial premises, and leading to the conclusion that simple theories have an increased chance of being true. It is perhaps particularly worthy of note that in his recent *Ockham's Razors*, Elliot Sober makes no claim that simplicity is sufficient to establish scientific realist theses.[11] As far as the present author has been able to tell, the only area in which there has been real progress in developing a justification of simplicity in this sense has been in the specific, restricted area of curve-fitting. But, as we have noted, this is not enough to help us make a case for scientific realism. I conclude that we have yet to see how we might establish realism via IBE, where the best theory is construed as the *simplest* theory.

4.7 Criteria Other Than Simplicity

So far in our discussion we have assumed "best theory" means "simplest theory". But, of course, simplicity is not the only property of theories that has been seen as making a theory good. Amongst the other properties that have also been suggested are: testability, empirical content, explanatory power, symmetry and coherence with other theories.

We will not here give a detailed discussion of all properties of theories that have ever been suggested to be "good-making". Rather, it will merely be argued that the prospects for justifying IBE with these other properties seems to be no better, and may often be rather worse, than the prospects for doing this with simplicity.

First, it is clear that the testability of a theory does not furnish us with a reason for saying it is true or likely to be true. To say a theory is testable is merely to say that there are tests to which we can subject it: it does not entail that the theory would *pass* those tests. And even if it did, the question would remain: does the fact that the theory passes tests constitute a reason for saying it is true? Given the underdetermination of theory by actual data, the question would arise: what reason do we have for preferring one theory to the others that would pass the same tests? Similar remarks apply to empirical content. A number of theories will have the same empirical content: what reason will we have for regarding one of them as more likely than the others?

Explanatory power runs up against similar difficulties. To say a theory has great explanatory power is to say it can explain some set E of observable phenomena, where E is in some sense a *large* set. But, if theory is underdetermined by actual data, there will be a number of theories that can explain E. And so, the question arises: what reason do we have to believe T is true, rather than any of the other theories that can explain E? We find ourselves back at square one.

---

[11] See Elliot Sober *Ockham's Razors* (Cambridge University Press, 2015). On pp. 144–145 Sober does say that the Akaike Information Criterion may have some relevance for what the realist is trying to do, but he makes no claim scientific realism can be established in this way.

Symmetry is widely regarded as a desideratum of theories, especially in theoretical physics. But it is plain, I think, that many of the points we made above about simplicity also apply to symmetry.

Let us begin by giving a definition of symmetry, as that notion is often used in physics. The value of a property of a system is said to be symmetrical with respect to an operation if and only if the operation leaves the value of the property unchanged.[12] We can illustrate this definition with an example. Suppose mass is conserved. Then the mass of some closed system S will be the same at all points of time. The mass of the system S will be said to exhibit "symmetry" in the following sense: the value of the mass of S will remain unchanged under the operation of "translation over time".

Many of the points we made about simplicity also apply to symmetry. Let us continue with the law of the conservation of mass as an example of a hypothesis that exhibits symmetry. Suppose we have made a number of observations of the mass of a system S, and all the observations are, allowing for the possibility of observational error, consistent with the hypothesis that mass is conserved for S. The question then arises: Do these observations support the hypothesis that mass is in fact conserved for S – including for points of time we have not yet observed?

There may well be some sort of "justification of our preference for symmetry" along the same lines as the justifications of simple curves earlier mentioned.[13] But even if this is so, it would get us no further in establishing a case for scientific realism. To see this, let us consider whether an inference of the following form might be used to establish realism:

Premise: In some restricted range of cases, whenever system S has been subjected to some operation O, the magnitude of property P has remained unchanged.

Conclusion: The magnitude of property P in system S remains the same no matter how S is subjected to operation O.

It is clear that we are here confronted with same dilemma as we were with our discussion of simplicity with curve-fitting. Is the property P observable or not? If it is, then even if the inference just given is a good one, it only establishes the distribution of an observable property in S: it does nothing to make a case for the existence or behaviour of any unobservable, theoretical property. If P is not observable, then the conclusion of the inference might be a scientific-realist claim, but we remain in the dark concerning how we might have good reason to believe the premise. Either way, the notion of symmetry would not appear to furnish us with a route to realism.

It has been suggested that a theory is good if it coheres with other theories.[14] Presumably, this would only be considered a virtue of a theory if the other theories

[12] This definition comes from K. Mainzer *Symmetries of Nature* (Berlin: Walter de Gruyter, 1996).

[13] That such a justification can be given is defended in the present author's *Explaining Science's Success: How Scientific Knowledge Works* (Acumen, 2014).

[14] That coherence with other theories is a desideratum of scientific theories has been defended by, for example, Larry Laudan Progress and Its Problems (Routledge and Kegan Paul), 1977.

with which it cohered themselves had merit. It would, for example, not be a point in favour of a contemporary theory if it "cohered" with the theory of phlogiston. So, let us say: It is a point in favour of a theory that it coheres with other theories only if those other theories themselves are good. But then, of course, the question arises: *what makes* those other theories good? To say those other theories are good because they cohere with still other theories would not, by itself, appear to be satisfactory: we might construct a work of pure fiction each sentence of which coheres with some other sentence in that work of fiction and yet still be without any reason to believe the whole. If we are to have good reason to accept as true or close to the truth some scientific theory we would, it seems, need more than mere coherence with other assertions.

A natural suggestion to make here is that perhaps a theory is good if it coheres with other statements, where we have independent grounds of some sort for accepting those other statements – for example, perhaps, because they have been confirmed by observation. Then a fundamental problem arises: what grounds do we have for regarding coherence as a good-making property of a theory? In the present context we are especially concerned with whether relations of coherence would give us good reason to believe a theory about unobservables. Would the fact that a theory T about unobservables "coheres with" a theory T* about observables constitute a good reason for believing T? For it to constitute such a good reason, it must surely be the case that we have good reason to believe some conditional such as: "If a theory with the attributes of T* is true, then we have good reason to believe a theory with the attributes of T is true", where T* is a claim about observables and T is a claim about unobservables. But whether this can be so is precisely the question with which are currently concerned! What we want to know is whether, and how, some observable state of affairs constitutes good reason for an unobservable state of affairs. We are led back to square one. The prospects, therefore, for the notion of coherence with other theories providing us with a justification of realism would not seem promising.

There are of course a variety of other properties of theories that have been claimed to make a theory good. Amongst these are: conservatism, accuracy, fruitfulness and agreement with philosophical and metaphysical beliefs. But none of these seem very likely to be able to give us good grounds for realism. To say that a theory is conservative is to say its adoption requires minimal change to the rest of our "web of beliefs". But that would only seem to give us reason to believe the theory if we were already in possession of good reason to believe the remainder of our web of beliefs was true. To be sure, it is always desirable for a theory to be accurate, but a failure of accuracy, by itself, would only appear to furnish us with reason to not believe the theory. More argumentation would be required to show that a theory about *unobservables* that was empirically *accurate* was thereby worthy of rational acceptance as a description of what is going on at the level of the unobservable. A theory is fruitful if it turns out to have unexpectedly high explanatory power, but one of the questions with which we are concerned is: what, if anything, entitles us to assert that a theory that can explain a lot is true in both its observational *and* theoretical, non-observational claims. And finally, agreement with philosophical and metaphysical beliefs would only seem to constitute good reason to accept a theory if we are already in possession of good reasons for those other beliefs, and our understanding of what makes a philosophical or metaphysical belief good is surely at a less developed level than our understanding of what makes a scientific theory good.

In conclusion, the prospects for developing an account of reasons for realism using properties other than simplicity would not seem to be very promising.

4.8 Lipton's Defence of IBE

One influential account of IBE is given in Peter Lipton's *Inference to the Best Explanation*.[15] Lipton's account is further developed in some of his subsequent writings.[16] However, as Lipton himself freely acknowledges, his approach does not furnish us with a satisfactory route to Scientific Realism.[17]

Let us begin by briefly outlining Lipton's account. For Lipton, to assert that IBE is a good form of inference is, in his terminology, to assert that *ceteris paribus* the "loveliness" of an explanation increases the likelihood that it is (at least close to being) true.[18] *What it is* for an explanation to be "lovely" is not truth, or closeness to the truth, or probable truth, but rather how much it would increase our understanding if it were true. As a rough first approximation, "loveliness" is "potential explanatoriness".

Lipton makes no claim that if an explanation that postulates unobservables is "lovely", then it is more likely to be true *in a way that would give support to scientific realism*; in fact, he explicitly denies he has any demonstration it does. But it is worthwhile briefly reminding ourselves why loveliness does not give us a route to realism.

For Lipton, the following attributes can increase the loveliness of an explanation: simplicity, unification, scope, precision and the giving of a mechanism.[19] And so, for Lipton, the thesis that IBE is a good form of inference entails the thesis that *ceteris paribus* these attributes increase the likelihood that an explanation is true or close to the truth.

We will not here consider whether the attributes listed by Lipton really do increase the explanatoriness of a theory. Neither will we undertake task of determining whether under some range of circumstances explanations that have these attributes have an increased likelihood of being true. We will restrict ourselves to the

[15] Peter Lipton *Inference to the Best Explanation* (Routledge, 1991).

[16] See for example Peter Lipton "Is Explanation a Guide to Inference? A Reply to Wesley C. Salmon" in G. Hon and S. S Rakover (eds) *Explanation*: *Theoretical Approaches and Applications*, (Kluwer Academic Publishers, 2001), pp. 93–120.

[17] See for example Lipton (1991), pp. 158–184, pp. 188–189.

[18] See Lipton (1991), esp. pp. 59–60.

[19] See Lipton (1991), p. 122.

matter of whether they have an increased chance of being true when the explanation in question is *about unobservables of sort that concerns the Scientific Realist*. It is, I think, clear that the properties cited by Lipton seem no more likely to lead us to the truth about unobservables than do the properties discussed in the previous section.

In Sect. 4.6 we noted that simplicity by itself did not seem to be sufficient to supply us with a route to realism to unobservables. There were notions of simplicity and related notions that did provide us with justifications of the thesis that simple theories were more likely to have continued predictive success. But the justification did not give us any reason to think that these predictively successful simple theories would also be true in the realist's sense of true. In particular, they did not give us any reason to think the unobservable entities postulated by the theories exist and behave as the theories says they behave.

Let us now consider unification. Sometimes postulating unobservables does enable us to give a unified explanation of what would otherwise be a disparate group of phenomena. For example, if we say that gases are composed of numerous massive objects too small to see, then we can explain the gas laws in the same way we explain, for example, the motions of the planets and things like pendulums and colliding billiard balls. But does the fact that we are in this way able to achieve a unified explanation mean that we are thereby justified in saying that the unobservable entities used in the explanation actually exist? People disagree on this question. While some philosophers hold that the unified explanation does give us good reason to say atoms and molecules exist, others – particularly those inclined to instrumentalism or some form of anti-realism – are inclined to deny that it gives us good reason. The question of whether the fact that an explanation is highly unified gives us good reason to believe in the existence of the unobservables postulated by the explanation is a question on which the supporters and opponents of Scientific Realism disagree. So: appeal to unification of explanation does not provide us with a *non-partisan* route to Scientific Realism.

Similar remarks apply to the *scope* of an explanation. An explanation with great scope applies to a wide range of phenomena. But, again, realists and non-realists disagree on the question of whether an explanation's having great scope constitutes a good reason for believing in the existence of the unobservables postulated by the explanation. Like unification, scope does not provide us with a non-partisan route to Scientific Realism.

Lipton says that the more *precise* the predictions made by an explanation, the lovelier it is. But does the precision of the predictions made by a theory increase our confidence that the explanation is true? Of course, there is a respect in which more precise predictions can provide us with more impressive confirmation of a theory than can vague or imprecise predictions. If a prediction is very precise, it is *a priori* less likely that it should, merely by lucky chance, be observationally confirmed. And the less likely it is that a confirmation should be merely due to lucky chance, the more likely it is that any *actual* confirmation should *not* be merely due to lucky chance.

But, of course, it is one thing to say that the subsequent confirmation of the predictions of a theory is not merely due to chance, and quite another to say that it is due to the truth of the theory, where "truth" is given a realist interpretation. On the face it, it would seem to also be a possibility that, although the theory gets things right at the observable, level it gets them wrong at the theoretical, non-observational level. Does the fact that the theory gets things right at the observational level constitute good reason for saying it also gets it right at the non-observational level? As noted in the previous section, this is a matter upon which realists and non-realists disagree.

Finally, Lipton says explanations are *ceteris paribus* better if they explain by means of a *mechanism*. Intuitively, it does seem to be plausible that a theory is more explanatory if it describes a *mechanism* bringing about the observed phenomena. But is this sufficient to justify realism? One difficulty is that there might be several possible mechanisms capable of explaining the phenomena. The thesis of the underdetermination of theory by data gives support to the idea that there will be many possible mechanisms. If there are several possible mechanisms capable of producing the phenomena to be explained, then we are confronted with the question: How are we to tell which of the possible mechanisms corresponds to what is actually going on at the theoretical level? It has been argued that none of simplicity, unity, scope and precision will do the job – at least in a way that would be acceptable to the sceptic about Scientific Realism. So, stipulating that the explaining theory must describe a mechanism does not give us the route to realism we require.

In summary, even if we grant that the attributes of theories specified by Lipton – simplicity, unity, scope, precision and the specification of a mechanism – increase the "loveliness" of an explanation, it has yet to be shown how an explanation that is lovely in these respects is more likely to be telling us the truth about the existence and behaviour of unobservable entities. To repeat, Lipton himself makes no claim they do, but it has perhaps been worthwhile getting clear ourselves why Lipton's account is not able to supply us with what we want.

Lipton provides a further defence of IBE in his discussion of the case of Ignaz Semmelweis's and the incidence of childbed fever amongst expectant mothers.[20] Semmelweis noted that childbed fever was much more common amongst expectant mothers who were examined by medical students than it was amongst those examined by trainee nurses. His explanation appealed to the fact that, prior to examining the expectant women, the medical students had been engaged in dissecting corpses. He hypothesised that the presence of "cadaverous matter" on the hands of the students was infecting the women. Semmelweis tested this hypothesis by getting the students to wash their hands with antiseptic after performing dissections. This resulted in an immediate decrease in the incidence of childbed fever amongst the women the students inspected.

Lipton argues that Semmelweis accepted the suggestion that the infections were due to cadaverous matter on the hands of the students because it was the *best explanation* of the observed facts. He sees Semmelweis as here using IBE.

From our point of view, the example is of interest because it seems to be an inference to the existence of something that *could not be observed with the naked eye*. The "cadaverous matter" on the hands of the students was, presumably, not visible

[20] See Lipton (1991), pp. 80–90.

to the naked eye. If we have here a case of IBE to something unobservable, perhaps then IBE *can* furnish us with a satisfactory justification of Scientific Realism in at least some cases.

It will be argued however that this case would not seem to lend a great deal of support to Scientific Realism.[21] This becomes apparent when we consider what Semmelweis did *not* have to say about the cadaverous matter. The only attributes his theory attributed to it were a disposition to cause childbed fever and a tendency to be removed by hand washing. His theory did not, for example, say that the cadaverous matter consists of numerous organisms too small to see: Pasteur's germ theory of disease did not become established until after Semmelweis's death. Semmelweis was unable to provide any *theoretical* account of how the material the trainee doctors got on to their hands during autopsy caused disease in the expectant mothers. His inability to give such an account was a factor in the difficulty he experienced in getting his ideas accepted. The only *explanatory* property his theory attributed to the cadaverous matter was its tendency to cause childbed fever. But this falls rather short of what we usually want Scientific Realism to provide. Consider, for example, the theory of atoms. This case considered in more detail in a later chapter, but the theory of atoms is more than just the theory that there are some unobservable "somethings" that are responsible, for example, for the behaviour of gases. It was (at least in an early form) the theory that the things are responsible for the behaviour of gases are masses, too small to see, that are moving around according to Newton's laws of motion. Duhem claimed that we are not entitled to say that atoms in this sense exist. He asserted that that perhaps all we are entitled to say is that there exist "somethings" that have the propensity to cause gases to behave in the way that they do. But Duhem is regarded as an *opponent* of realism about atoms. His position, however it is to be classified, is not a form of Scientific Realism. More generally, to merely assert that there exists "something" that has a tendency to produce events at the macro-level, without specifying the nature of those "somethings" would not seem to count as a form of realism. But this is all that Semmelweis's theory does with respect to "cadaverous matter". It simply tells us there is something that causes childbed fever, and that it can be removed by washing the hands. As such, Semmelweis's theory would not seem to count as an example of the type of theory in which Scientific Realists are particularly interested. Perhaps we can arrive at Semmelweis's theory by IBE, but this would not seem to give much support to Scientific Realism.

---

[21] One point to note is that although the "cadaverous matter" was not visible, it was not altogether undetectable by the unaided senses. It had a putrid smell, and Semmelweis recommended washing the hands with calcium hypochlorite because this proved to be the most effective method of removing the smell. So, Semmelweis's explanation did not postulate something *entirely* unobservable to the unaided senses. But Scientific Realists surely want to be able to say we can have good reason for postulating the existence of something that is not detectable to *any* of the unaided senses. However, this point is perhaps is of no great relevance in this context since, plausibly, Semmelweis's explanation would still have been a good one even if the cadaverous matter had not had a detectable smell.

4.9 Kitcher's Galilean Strategy for Defending IBE

In his "Real Realism: The Galilean Strategy",[22] Philip Kitcher develops a defence of the idea that the success of a theory does give us good reason to believe it is true. In articulating this approach, Kitcher refers to an argumentative strategy used by Galileo in defending his use of the telescope. Galileo's critics asked what reason there was to believe the telescope was giving us an accurate representation of, for example, the surface of the Moon. Galileo's reply was that we could verify that the telescope gave us accurate representations of features of objects on the surface of the Earth too distant to make out with the naked eye. For example, suppose that the telescope told us that a distant tower had three windows, where this feature was not visible to the naked eye. We could, nonetheless, verify that it had three windows by travelling closer to the tower and seeing, with our unaided senses, that it did in fact have three windows. In this way we could get independent confirmation that the telescope was giving us an accurate representation of an object which, at the time of viewing, could not be seen with the naked eye. Galileo argued that if the telescope was giving us an accurate representation of an object such as a distant tower, it seemed reasonable to believe it was also giving us an accurate representation of even more distant objects, such as the Moon or planets.

The argumentative strategy that Galileo here used can perhaps be represented as follows. (1) Establish that some technique reliably gives us accurate representations at the directly accessible (e.g. terrestrial) level. (2) Make a case for the thesis that if it reliably gives us accurate representations at the directly accessible level it will also give us accurate representations at less directly accessible (e.g. non-terrestrial) levels. (3) Draw the conclusion that the technique *will* also give us accurate representations at less directly accessible levels. Kitcher refers to this as the "Galilean Strategy".

Kitcher suggests the Galilean Strategy can supply us with an argument for Scientific Realism. The way we typically go about forming beliefs about our immediate environment – the environment containing tables and chairs and the like – on the whole does seem to give us true or accurate beliefs about those things. But, just as the distance of a tower, or an object like the Moon, does not affect the reliability of a telescope, so the size, or relative lack of accessibility of the entities to which we might apply our typical methods of forming beliefs ought not affect the reliability of those methods. And so, we may conclude, if we extend the methods we typically use to arrive at beliefs about things like tables and chairs to the formation of beliefs about, say, atoms, then those methods still ought to give us truths. We may therefore more generally assert that our beliefs about the observationally less-accessible things postulated by science are probably true, and hence that we are entitled to say Scientific Realism is correct.

[22] See P. Kitcher "Real Realism: The Galilean Strategy" in *The Philosophical Review*, 110, (2001), pp. 151–197.

It might perhaps be suggested that Kitcher's "Galilean Strategy" could be used to supply us with a justification of IBE as a route to Scientific Realism. After all, there seem to be many situations in which IBE does appear to work very well as a method for giving us truths about observable things. For example, detectives successfully use something that looks rather like IBE to determine who committed a crime.[23] Historians and archaeologists frequently use it in interpreting some event or artefact from the past. And we all use it in innumerable ways in everyday life: we might, for example, say a mouse is responsible for left-out food disappearing. When it comes to reasoning about the behaviour of observable objects, IBE seems to work very well. So, the question arises: why should it not also work well when we come to apply it to entities that are too small to see? In fact, it would seem that *an Eddington inference* leads us to expect that IBE ought to work when we come to apply it to the unobservable realm.[24] The Eddington inference is as follows:

Premise (1): IBE has reliably led us to true conclusions when it has been applied to states of affairs in the observable realm.
Premise (2): It would be an improbable fluke if the region of applicability of IBE happened to coincide with the region we can perceive with the unaided senses.

Conclusion: IBE will probably reliably lead us to true conclusions when we come to apply it to regions that cannot be perceived with the unaided senses.

Here it will be argued that the above argument does *not* successfully defend the applicability of IBE to the unobservable realm. More specifically, it will be argued that Premise (1) is false.

There is a qualification that needs to be made here. Although it will be argued that Premise (1) – at least in a form required for it to act as a justification of IBE – is false, it will not be claimed that our inferentially arrived at beliefs about observable things are false. Rather, it will be argued that, strictly speaking, we do not arrive at them by IBE *simpliciter*. It will also be argued that the means by which we *do* inferentially arrive at conclusions about *observables* is in fact rather similar to the way in which, on the view adopted here, we inferentially arrive at conclusions about unobservables. However, it is not possible to argue for this last thesis in detail until we have given a full exposition of the method advocated here. We do not do this until the next chapter: so we must postpone a full discussion of this matter until then. However, here a case

[23] In *Philosophy and Scientific Realism*, p. 47, J. J. C. Smart notes that the way in which a detective might put various pieces of evidence together to arrive at the hypothesis that a burglar was responsible for a crime bears some similarity to the way a scientist forms hypotheses. Smart's passage is sometimes seen as an early statement of the "no miracles" argument. But Smart can perhaps also be seen as getting as something similar to Kitcher's Galilean strategy. If the detective's activities have led to the hypothesis that there is a burglar, and it turns out that there really was a burglar, the we have evidence that the detective's methods have led us to truths. And if the methods of the scientists are similar to those of the detective, then would thereby have reason to believe the methods of the scientist also leads us to truths.

[24] That this defence of IBE can be represented as an Eddington inference was pointed out to me by a referee for this series.

will be made for the thesis that, in inferentially arriving at conclusions about observables, there is a sense in which we do not use IBE *simpliciter*.

Let us consider a specific type of case in which, it seems, IBE has had a pretty good track record of leading us to true conclusions. The type of case we will consider is a detective observing the circumstances surrounding a crime scene. Assume a detective is confronted with the following observations: a window shows signs of having been opened by force, the window is on the second storey, a drain pipe leads from ground level up to the window, the window opens on to a room and a case in the room has been opened and valuables are missing. A plausible hypothesis here is that the burglar climbed up the drain pipe, forced open the window and took the valuables from the case.

This hypothesis might seem pretty obvious, but it is worth noting a number of things that the detective has here *assumed*. One piece of evidence is that the window shows signs of having been forced open. Perhaps it is cracked and splintered near the latch. So, it seems reasonable to hypothesise that there was some sort of agent outside the window, causing this cracking and splintering. But the detective, very naturally, *assumes* that this agent was a human being. This same assumption is made when the detective hypothesises that the agent responsible for the cracking and splintering got in to a position outside the window by climbing up the drain pipe, rather than, say flying like a bird.

There are innumerably many possible hypotheses that would explain how the window was forced open. Perhaps a human being did it, but perhaps also a levitating gremlin did it, or some hitherto unknown species of intelligent bird, or a self-organising colony of rats, and so on. It is not being suggested that perhaps some of these alternative hypotheses might be better than the hypothesis that a human did it. The hypothesis that a human being did it is obviously "better" than any of the others just suggested, in the sense of being *more worthy of rational belief than the others*. But let us consider just *why* it is more worthy of rational belief.

The short answer, surely, is because we already know that human beings exist. A slightly lengthier answer is: We already know that there exist human beings and that they have the power to climb up drain pipes but not to levitate or fly like a bird. And they have the power to force windows open. But we do not know that gremlins exist, neither do we know of birds intelligent enough to force windows open, or of self-organising colonies of rats. The detective (presumably) has not observed the *particular* human being that climbed up the drain pipe and forced open the window, but he has observed *other* human beings that have the power to do similar things. However, he has not observed birds with the power to force open windows, and he has not observed gremlins at all. And something like this seems to be the reason why we regard the hypothesis that a human being is responsible as being more worthy of rational acceptance than the other hypotheses.

If this is right, then there seems to be a sense in which, although the explanation that says a human being did it *is* the best, it is not the best purely in virtue of its own "*internal* merits". It is not the best (merely) in virtue of being, say, simpler than the others. It is, rather, the best at least in part because we know, by observation, that that type of entity it postulates does exist "anyway". Adopting this explanation does not require us to postulate the existence of any never-observed *type* of entity. The fact that we have observed humans that can climb drain pipes, force open windows and so on lends credibility to the idea that in this case a human was responsible even though we did not actually observe what took place.

In the next chapter it will be argued that there are at least some cases from the history of science in which Eddington inferences support explanations that postulate unobservables in something the same way that prior observations of human-beings and their powers support the hypothesis that, in the case under consideration, a human-being was responsible for the robbery. Moreover, it will be argued that it is only because these explanations postulating unobservables are supported by Eddington inferences that those explanations are worthy of rational acceptance.

In the light of this, let us consider again the significance of the claim that IBE has worked well when applied to observable objects. Perhaps it has, but when we apply IBE to observable objects the resulting explanations are *also* supported, in a special way, by observations. They are supported by observations of objects of the same sort, and which have the same causal powers, as those entities that figure in the explanation. This procedure has worked well, but the fact that this procedure has worked well would not seem to give us any reason to think IBE would work well if it led us to postulate the existence of entities that *do not* have the properties or powers of those we have already observed.

We can perhaps bring this out if we tried to imagine a hypothetical use of IBE to observable clues or data, but which did not postulate entities we already knew to exist. The detective could, perhaps, explain how the window came to be open by postulating a levitating gremlin that telekinetically forced the window to open. Explanations postulating levitating gremlins as being responsible for the clues at a crime scene would presumably *not* turn out to have a good track record. And, of course, we regard the explanation that says a human did it as clearly rationally preferable: it is rationally preferable because it only requires us to postulate an entity of the same sort and with the same powers as entities we already know exist.

In summary, there is a respect in which IBE has worked well at the observational level. But in the cases in which we are confident it has worked well, it has not postulated entities that are unlike those we already know by observation to exist. It rather only postulates entities that have the same powers as those we already know by observation to exist. So, it is not IBE "by itself" that has worked well at the observational level, it is rather IBE supported in a particular way by prior observations of entities similar to those used in the explanation. In this respect, the supposed or alleged cases of IBE working well at the observational level are actually more like examples of scientific explanations that are also supported by Eddington inferences. This matter is discussed further in the next chapter. But the point to note here is that the good track record of detectives working out what has taken place at a crime scene does not give us good reason to think that IBE, by itself, will reliably lead us to truths about unobservable entities.

Let us now summarise the main points of this section. We have been considering whether Kitcher's "Galilean Strategy" can be used to show IBE will work when leading to conclusions about unobservable entities. It has been argued that this suggested way of justifying realism is not successful. The reason why it is not successful, it has been argued, is because the method that has worked well at the observational level is not IBE *simpliciter*.

4.10 Novel Predictive Success

We are considering the idea that we might be able to establish scientific realism by means of inference to the best explanation. But, of course, this leaves unanswered the question: inference to the best explanation *of what*? So far, we have (mostly) been assuming that what is to be explained is the predictive success of science. But it is customary to draw a distinction between two types of predictive success. These may be called novel predictive success and non-novel, or familiar, predictive success. A number of authors have suggested that the case for realism with respect to a theory is stronger when that theory has novel predictive success.[25] So, might we perhaps say that realism is to be justified on the grounds that it provides *the best explanation of the novel predictive success of science*?

This type of argument for realism can be given a "material mode" formulation or a meta-level "formal mode" formulation. In its "material mode" formulation it says we are entitled to accept as true a particular scientific theory T if that theory is the best explanation of some *novel* phenomenon N. The "best explanation" is the theory T itself, and the explanandum is the novel phenomenon N. In its "formal mode" formulation, the best explanation is not the theory T itself but rather the hypothesis that T is true or close to the truth, and the explanandum is the novel predictive success of science.

Let us begin by giving a more explicit statement of the "material mode" form of the argument: if theory T successfully predicts novel phenomenon N we are entitled to accept that T is true. An immediate difficulty arises for this suggestion: Suppose some alternative T* also successfully predicted the same novel phenomenon N. What, if anything, would entitle us to accept T in preference to alternative T*? If we say T is to be accepted rather than T* because it is better, we are back to our original question: What entitles us to accept IBE?

It might perhaps be protested that in the above argument it is assumed that there will be a *number* of different theories T_1, …., T_n that are all capable of explaining some novel phenomenon. But perhaps this will not be so: perhaps there will be only one way of explaining some particular novel phenomenon N.

However, it will be argued that even if in some cases there is only one explanation T of some novel phenomenon, still, we cannot avoid appealing to something at

[25] Authors who have emphasised the importance of novel predictive success include Alan Musgrave "The Ultimate Argument for Scientific Realism" in Robert Nola (ed) *Relativism and Realism in Science*, Kluwer Academic Publishers, pp. 229–252. Jarrett Leplin *A Novel Defence of Scientific Realism* Oxford University Press, 1997, Stathis Psillos *Scientific Realism: How Science Tracks the Truth*, Routledge 1999.

least closely resembling IBE if we are to establish realism with respect to T. Suppose there was one, and only one, theory T that successfully made some novel prediction. Then, even so, we might be reluctant to adopt a realist view of T if it was a *very bad* explanation.[26] This seems to show that, even if some theory is *the only* theory that makes some novel prediction, it must be somehow "sufficiently good" if it is to be worthy of rational acceptance. But we are now confronted with questions at least closely resembling those that troubled us earlier: How good must a theory be to be worthy of rational acceptance? And, if it is good enough, what are the properties it has in virtue of which it is worthy of rational acceptance? We are at least very close to being back to our original question.

Jarrett Leplin has developed a defence of the argument for realism from novel predictive success.[27] A key aspect of Leplin's account is his *definition* of "novel predictive success". According to Leplin, in order for an observation to count as a novel observational result O of a theory, it must satisfy two conditions. The first of these is an "independence condition", the nature of which need not concern us here. The second is a "uniqueness condition", which may be expressed as follows:

> *Uniqueness*: There is some qualitative generalisation of O that T explains and predicts, and of which, at the time that T first does so, no alternative theory provides a viable reason to expect instances.[28]

Roughly, we may say that in Leplin's sense, a theory T has novel predictive success, at some point in time *t*, if and only if (provided it also meets the independence condition) it successfully predicts O and no theory around or known at *t* also does so. It is important to note that on Leplin's view, for a theory to meet his uniqueness condition, it is not necessary that it be, in some Platonic sense, the only existent theory that predicts O, merely that it is the only theory *known by scientists at that time* that does so. But: is the fact that a theory T is "unique", in Leplin's sense of uniqueness, sufficient to make it rational to believe T? For all we know, it might be possible to come up with alternative explanations of, or predictors of, O. We could certainly do so if we are prepared to allow highly *ad hoc* theories.[29] And, as our discussion in Chap. 3 showed, the history of science would suggest there seems to be a pretty good chance that a theory postulating unobservable entities different from those of T might be discovered in the future. If such alternatives are "out there", or seem reasonably likely to be so, what makes it rational to believe T? It might be claimed T is the *best* of the alternatives. But then we are confronted once again with our original question: what, if anything, justifies us saying the best theory is worthy of rational belief? And: if we allow that better explanations might be discovered in the future, it is not even clear T *is* the best. So: even if T satisfies Leplin's uniqueness

---

[26] This point is made in Musgrave "The Ultimate Argument for Scientific Realism" in Nola.

[27] See Jarrett Leplin *A Novel Defense of Scientific Realism* (Oxford University Press, 1997).

[28] See Leplin *op cit*., p. 77.

[29] Leplin seems to allow this in his discussion of underdetermination and empirical equivalence. See Leplin, *op cit*., p. 155.

condition as an explanation of O at a particular point in time, it seems more work needs to be done to show it is thereby worthy of rational acceptance.

So, in summary, appealing to the notion of novel predictive success would not appear to quite get us there in our quest for a good argument for realism.

It might be suggested that the notion of novel success might yet enable us to establish *some* form of realism, such as structural realism, even if could not quite get us to "full-blown" realism. This is discussed below.

We have been considering the "material mode" version of the argument for realism from the notion of novel predictive success. But, as noted above, there is also a "meta-level", formal mode version of the same argument. In the "formal mode" version of the argument, the explanans is the assertion that some theory, or theories are true, or close to the truth, while the explanandum is the novel predictive success of those theories. But it is clear that much the same difficulties are confronted by this "formal mode" version. Suppose it is true that the hypothesis that certain theories are true provides us with the best explanation of novel success. The question arises: What reason, if any, do we have for believing that the fact that ""Theory T is true or close to the truth" is the *best explanation* for the novel success of T." does in fact constitute good reason for accepting "Theory T is true or close to the truth"? In fact, the situation here is arguably even worse than it was with the "material mode" version of the argument. Theory T is a *scientific* theory, and we have a reasonably clear idea of what it is for a scientific theory to constitute an explanation of some phenomenon, and we also seem to have a rough idea of what makes one theory better than another. But the assertion "Theory T is true or close to the truth" is not itself a *scientific* theory. It is perhaps a meta-scientific, or perhaps philosophical theory. It is perhaps not so clear what makes this theory better than alternatives; it is also perhaps not so clear what, if anything, would entitle us to say that the fact that it is the best means it is true. And there is also, of course, the question of whether the notion of "truth" is itself an *explanatory* concept. It has been argued it is not.[30] So: for all these reasons, the meta-level "formal mode" version of the argument would seem to be confronted with rather more obstacles than the "material mode" version. And the obstacles confronting both versions are considerable.

---

[30] See for example, John McDowell "Physicalism and Primitive Denotation: Field on Tarski" in *Erkenntnis*, 1978, v.13, pp. 131–152. and Michael Levin "What kind of explanation is truth?" in J. Leplin *Scientific Realism* (Berkeley, University of California Press, 1984), pp. 124–139.

4.11 Deployment Realism

A further refinement to IBE is offered by "deployment realism". This has been developed in particular by Stathis Psillos.[31] According to Psillos, we are entitled to continue to accept those parts of a theory that are "responsible for its empirical success", even if other parts of the theory have been refuted.[32]

In this section, the general idea of deployment realism will be discussed first; more specific versions discussed later.

One example from the history of science that seems to be a natural candidate for deployment realism is Fresnel's theory of light. As we have noted, this theory successfully predicted the "Poisson spot" in the middle of a perfectly round shadow. Laudan, and others, have pointed out that Fresnel's theory was an ether theory of light, and as such was false.[33] Laudan sees this as supporting the sceptical thesis that the empirical success of a theory does not justify us in adopting a realist stance towards that theory. But this case seems to comport very well with "deployment realism". Intuitively, there seems to be a clear sense in which the ether does not *play a direct*, *contributory role* in deriving the prediction of the white spot: it is not "deployed" in its derivation.[34] What is deployed in the derivation of the prediction is the particular transverse wave character that Fresnel attributes to light. And so, deployment realism suggests we are entitled to attribute this particular wave character to light.

One question naturally arises: How well does deployment realism fare against the pessimistic meta-induction? Does the history of science show that even the mechanisms "directly deployed" in deriving successful novel predictions have subsequently turned out to not exist? It has been argued by Timothy Lyons that deployment realism does not fare too well, and there are cases in which the mechanisms deployed have subsequently been found not to exist.[35]

We will not here enter in the debate about whether or not deployment realism is vulnerable to the pessimistic meta-induction.[36] Instead, it will be argued that there

---

[31] Deployment Realism is developed in Stathis Psillos *Scientific Realism*: *How Science Tracks the Truth* (Routledge, 1999). A form of deployment realism is also advocated in Wright, J. *Science and the Theory of Rationality* (Avebury, 1991).

[32] Psillos, *op cit*, p. 108.

[33] See Larry Laudan "A Confutation of Convergent Realism" in Jarrett Leplin (ed), *Scientific Realism*, (University of California Press, 1984), pp. 218–249., esp. p. 225.

[34] This point is made in Philip Kitcher *The Advancement of Science* (Oxford University Press, 1993) pp. 144–149.

[35] See T. Lyons "Scientific Realism and the Pessimistic Meta-Modus Tollens" in *Recent Themes in the Philosophy of Science* edited by S. Clarke and T. Lyons (Kluwer Academic Publishers, 2002), pp. 63–90. See also G. Doppelt "Empirical Success or Explanatory Success: What Does Current Scientific Realism Need to Explain?" *Philosophy of Science* **72** (2005), pp. 1076–1087.

[36] See, for example, Doppelt *op cit*, also Doppelt, G. "Reconstructing Scientific Realism to Rebut the Pessimistic Meta-Induction" in *Philosophy of Science*, 74, (2007) pp. 96–118. In the opinion of the present author, deployment realism would seem to get in to difficulty with cases such as the caloric theory of heat, phlogiston and Rankine's theory of heat. For an opposing point of view, see

is a different, and perhaps more fundamental, difficulty for deployment realism. It will be argued that deployment realism has the same shortcoming as other approaches to realism that rely on IBE. It assumes that we are entitled to say that our best theory – where this perhaps is something like our simplest, or most plausible, or most symmetrical, etc. theory – is worthy of rational acceptance. But we have yet to see what entitles us to say that this is so. It will further be argued that, once this is noted, it becomes apparent that, as a way of identifying those parts of a theory towards which realism is warranted, the criterion of "deployment" is somewhat misdirected or off-target.

The difficulty can be stated quite briefly. Suppose some property or mechanism or entity P plays a direct, contributory role in theory T's explanation of some novel phenomenon N. Then: if, generally, our theories are underdetermined by the data on which they are actually based, there will be some alternative explanation T*of N. Alternative theory T* might not postulate P at all: it might essentially deploy some alternative mechanism P*. And so, the question arises: what grounds would we have for accepting T rather than T*? If it is asserted that we have good reason to accept T because T is better, then we are back to our original question: What, if anything, entitles us to inference to the best explanation?

This point can perhaps be brought out more vividly by considering a simple, hypothetical example. Suppose the observations some scientist wishes to explain are those of the motion of the hands of a clock around a clock face. One hypothesis advanced to explain these observations is that inside the clock is a mechanism consisting of cogs, springs and a pendulum. Now, suppose, more specifically, that in the version of this explanation as advanced by the scientist, the interior mechanism of the clock is postulated to contain a particular cog C, where:

(i) Cog C is asserted to have 60 spokes
(ii) Cog C is asserted to be made of copper.

We may further assume there is no reason available to the scientist why the cog *needs* to be made of copper.

The fact that the postulated cog is claimed to have 60 (rather than 59 or 61) spokes we will assume *is deployed* in the explanation of the observed motions of the clock-hands: it maybe explains why the second hand takes 60 (rather than 59 or 61) "clicks" to complete a full circle. But the fact the cog is claimed to be made of copper is not deployed. Presumably, the explanations of the motions of the hands would still go through if the cog were made of iron, or chromium, or many other metals.

But still, none of this need give us good reason for saying that cog C really does exist inside the clock. There may be some altogether different way of explaining the motion of the hands. Perhaps the motions of the hands could be explained by saying that inside the clock there is a mechanism consisting of a battery, magnets, an electric motor and so on. In this mechanism, there is no 60-spoked cog at all.

M. Alai "Deployment vs Discriminatory Realism" in *New Thinking about Scientific Realism*, http://www.philsci-archive.pitt.edu/10551/

Does the scientist have good reason to believe in the existence of a 60-spoked cog inside the clock? The answer, surely, is that the scientist has good reason to believe this only if she is already in possession of good reason to believe that some version of the first hypothesis (that is, the hypothesis postulating cogs, pendulums and so on) is true rather than the alternative mechanism involving the electric motor etc. But, there is nothing in the approach of "deployment realism" that explains how the scientist might have good reason for this belief.

Assuming that both explanations can account for the empirically observed motions of the clock hands, we would, it seems (at least if we are to rely on IBE) need to appeal to some other (non-empirical) criteria, such as simplicity, to give us reason to prefer one explanation to the other. But we have yet to see how such criteria (such as simplicity etc.) increase the probability of truth.[37]

On this view, the notion of deployment is not quite what is required to give us good reason to believe realism. What the notion of deployment does, rather, is identify those aspects of a postulated mechanism that have a contributory role in producing an effect *of the specific mechanism*. It differentiates between the causally contributory and causally idle aspects of a specific, described mechanism. The 60 spokes of cog C have such a contributory role; the (claimed) fact it is made out of copper does not. But that is not enough for us to be justified in saying that a cog with 60 spokes *exists*. We also need good reason for saying that the explanation employing the cog, rather than the explanation using the electric motor etc., is the explanation that is *true*. But what grounds do we have for saying this explanation is true? If we say it is *true* because it is *best* we are back to square one: what grounds do we have for IBE?

Psillos has argued that the case for deployment realism is especially strong if the deployed mechanism yields novel predictions that turn out to be correct.[38] In our hypothetical example of the clock, neither theory of its internal mechanism seems likely to lead to what we would call a novel prediction. But it might perhaps be suggested that deployment realism becomes a viable position if it is stipulated that the prediction to which the deployed mechanism leads is a *novel* prediction. However, it is not too difficult to see that the same difficulty arises. If two or more theories, each mentioning two different essentially deployed mechanisms, both led to the same novel prediction N, then we would be back to our original problem: what reason do we have for preferring one of the theories to the other?

It is worth noting here that the history of science furnishes us with some examples of (by our lights) *false* deployed mechanisms that yielded novel predictive success. Two examples are the phlogiston theory of combustion and Rankine's theory of heat. The phlogiston theory of combustion had at least one novel predictive

[37] Psillos op *cit*, appeals to simplicity, coherence, consilience and related criteria to here justify a preference for one theory over another.

[38] See Psillos *op cit*, pp. 107–108. This position is also defended in Wright, J. *Science and the Theory of Rationality* (Avebury, 1991).

success.[39] And Rankine's theory had numerous predictive successes.[40] But both theories are, we now believe, false. Currently accepted theories (specifically, the oxygen theory of combustion and the kinetic theory of heat) offer alternative ways of explaining the novel phenomena predicted by phlogiston theory and by Rankine's theory. So, in these cases we can assert that the data available at the time that phlogiston theory was advanced, and the data available at the time Rankine's theory was advanced, could have also have been explained by other theories. They could also have been explained by the theories *we* now believe to be true.[41] This illustrates how even deployed mechanisms, that yield novel and subsequently confirmed predictions, are underdetermined by the observations on which they are based. And if the theory of those deployed mechanisms is underdetermined by the data, we are confronted with the question: "What entitles us to believe that the theory that postulates those deployed mechanisms *is true*?" If we say we are entitled to believe it is true because it is best, we once again find ourselves back to our original question.

There is a refinement to deployment realism that ought to be considered here. Let us return to our example of the theories about the internal workings of a clock, specifically, to the theory that the internal workings consisted of cogs, springs, pendulums and so on. We noted that the fact that cog C was postulated to have 60 spokes contributed to the predictions the theory made about the motions of the clock hands, while the specific substance out of which the cog was made did not. We might further explicate this in terms of the propensity of an alteration to the internal mechanism of the clock to *make a difference*. Altering the number of spokes on cog C would make a difference to the empirical predictions of this particular theory concerning the way the hands would move. But changing the postulated make-up of the cog from, say, copper to iron presumably would not. This suggests a way of explicating the idea that a particular entity or structure or property is deployed in the derivation of a (novel) prediction:

Let C be a component of a mechanism M postulated by a theory T, and let P be a property of M. Then, property P *makes a difference* to the (novel) empirical predictions of theory T if and only if changing the value of P, while leaving all other

---

[39] One novel predictive success of phlogiston was its prediction that heating the calx of mercury would result in the creation of an "air" that supported combustion more vigorously than does ordinary air.

[40] See Hutchison (2002), pp. 94–95.

[41] A reader of an earlier draft of this chapter objected that it is highly implausible to say that, for example, the phenomena available at the time Rankine was working could have been explained by the theory we now believe to be true, since many of the concepts – specifically, quantum-theoretic concepts – were not available to any theorist working at the same time as Rankine. However, it seems to me that this objection relies on an ambiguity of the expression "could have been explained by". Of course, as a matter of practical fact, no one working at the time of Rankine could have come up with a modern quantum-mechanical explanation. But still, an explanation using modern concepts would *logically entail* descriptions of the phenomena with which Rankine is concerned. Rankine's phenomena are "explainable by" modern theory simply in the sense that (descriptions of) Rankine's phenomena are logically entailed by modern theory. The fact that no one at the time of Rankine would have been able to come up with our modern explanation is irrelevant.

aspects of M unchanged, results in a change in the (novel) empirical predictions made by T._____(4)

However, it is clear that, as a way of identifying the parts of T towards which we are justified in adopting a realist stance, (4) will not do. It gets us no further in explaining what *justifies* us in postulating a particular deployed mechanism. Consider again our competing theories of the internal workings of the clock. Changing the number of spokes on a cog would change the predictions of the theory, changing the colour of the cogs or the substance out of which they were made would not. But this does not entitle us to assert that there actually exists inside the clock a cog with 60 spokes unless we already have good reason to believe that some version of the cog-spring-pendulum theory, rather than the electric motor theory, is true.

4.12 Underdetermination Again

It is perhaps appropriate here to briefly re-iterate a point made in the previous chapter. The argument for deployment realism discussed in the previous section was, it was argued, still vulnerable to an objection: Even if a deployment realist were to restrict their realist-claims to those mechanisms that are directly deployed in the derivation of successful novel predictions, they would still have to deal with the fact there might be *many* such mechanisms capable of making the same novel predictions. They would, that is, be confronted with a version of the argument against realism from underdetermination. But, it might again be protested, underdetermination does not just create a difficulty for this form of deployment realism. Some arguments that might be given in support of the underdetermination thesis – such as, for example the idea we are living in a simulation – also seem to undermine common-sense realism about material objects. But *the view adopted here* accepts, and in fact relies upon, the thesis that common-sense realism about material objects is rationally justified. And so, it might be protested, relying on the underdetermination thesis as a way of showing that Eddington inferences are required to establish the truth of Scientific Realism might actually have the effect of showing the approach of Eddington inferences to also be rationally untenable.

However, we gave a reply to this difficulty in the previous chapter. The position adopted here is that we reply to arguments for scepticism about material objects by adopting a "Moorean" view of these matters. On the Moorean view, although there exist arguments for scepticism about material objects, it is more likely that there is something wrong with the arguments than that the sceptical conclusion is true. So, on such a view, in so far as arguments for underdetermination from simulation hypotheses and the like are incompatible with common-sense realism, it is more likely that there is something wrong with the arguments than that common-sense realism is false. Common-sense realism about tables and chairs and so on remains intact.

Adopting this Moorean view gives us what we require. Since common-sense realism remains intact, it is still the case that we *defend* Scientific Realism if we succeed in showing that a belief in the existence of unobservable objects is more or less as good as belief in observable objects. But also: the Moorean view only establishes realism about things like tables and chairs. It does not establish the truth of Scientific Realism. Some other set of arguments are required for this purpose. And it is argued here that Eddington inferences are capable of playing this role.

4.13 Reliabilism and the History of Science

In Sect. 4.11 it was argued that it is at best doubtful whether saying that deployed mechanisms that lead to novel predictive success exist is consistent with the history science. Phlogiston theory and Rankine's theory of heat contained reference to such mechanisms that subsequent investigations found to not exist. But there is also another point that can be made here. Let us suppose, for the sake of the argument that realism about deployed mechanisms that had novel success *did* square with the history of science. Would this mean that this type of realism was satisfactory? Not necessarily. Here, our concerns are primarily *epistemological*. We want an understanding of what, if anything, *justifies* a belief in unobservable entities. Even if the thesis "If a deployed mechanism X leads to novel predictive success, then X exists" has no known historical counter-examples, that need not *explain what makes belief in X rationally justified.* We can bring this out by considering another – admittedly very far-fetched – example. Suppose we discovered that any theory advanced immediately after a scientist had consumed a large amount of coffee was never refuted by later testing. Then we might suggest the following:

If a theory T has been advanced after the scientist suggesting T had consumed a large amount of coffee, then any unobservables postulated by T exist.____(5)

We are assuming (5) would "square with the history of science" in the sense that the history of science did not furnish us any counter-examples to it. But even given this unlikely historical fact, I don't think we would say (5) gave us a satisfactory account of that in virtue of which *belief* in realism is rational. It would fail for the same reasons that, it was argued in Chap. 2, reliabilism fails to give us a satisfactory account of that in virtue of which a belief is rational.

4.14 The Argument from Concordance, or the Agreement of Independent Methods

One important variant on the idea that realism provides the best explanation of the success of science is that, in some cases, realism provides the best explanation of "concordance", or the agreement of independent methods. It is, perhaps, this agreement that Putnam and others have in mind when they say that, without realism, the success of science would be "a miracle". However, in this section it will be argued that the agreement between independent methods does not, by itself, provide us with a satisfactory justification of realism.

Sometimes if two independent methods give us the same result then our confidence in the reliability of both methods is thereby increased. Suppose we wish to determine the height of mountain. One way of doing this has a climber with a measuring rod, laboriously measuring how much higher they get with each step. The other might use radio signals from a satellite. We might have doubts about the reliability of both methods. We might doubt the first method because we think it likely the climber will make a mistake. And we might have doubts about the second because, at the time of testing, the technology involved is new. But suppose both techniques give exactly the same result, telling us that the mountain is, say, 15, 873 feet high. In this case our confidence that the mountain really is 15, 873 feet high would be much higher if we had used only one method but got exactly the same result. And the reason for our confidence seems to be along the following lines: unless both techniques are accurate, and measured the height of the same mountain, it would be an extraordinary fluke that they give us the same value. It is surely highly unlikely that such a fluke should have occurred. So, we have good reason to believe both techniques are accurate and hence that the height of the mountain is 15, 873 feet.

The above example does, quite clearly, give us good reason to believe both methods are reliable. But it is not immediately clear how this relates to *realism*. More specifically, it is not entirely clear how the agreement between the two methods would make us more inclined to adopt "realism with respect to the height of the mountain".

There are at least two ways in which the truth of "The mountain is 15, 873 feet high" might be subject to sceptical doubt. We will refer to these as "sceptical doubt one" and "sceptical doubt two".

Sceptical Doubt One: It is doubted whether the mountain has a height of 15, 873 feet *rather than* 18, 875 feet.

Sceptical Doubt Two: It is doubted whether "The height of the mountain is 15, 873 feet" *is true as statement about the objective spatial properties of the mountain* (rather than merely telling us what result we would obtain by measurement operations).

There seems to be a sense in which the agreement between the two methods is, plausibly, a very good response to sceptical doubt one. But it is not so clear it is an appropriate response to sceptical doubt two. It is not so clear it is a response to someone who is not a realist about the property of spatial extension. In fact, considered as a possible response to Sceptical Doubt Two, it seems to be rather beside the point. The distance between the base of the mountain and its top might be 15, 873 *somethings*, but the nature of those "somethings" might be open to further question.

One reason why it is open to question is because sceptical doubt two implies, or suggests, its own alternative explanation for the agreement between the methods. On this alternative explanation, there is no such thing as the height of a mountain, considered as an objective property of a material thing extended in physical space. What there is instead, playing the same explanatory role, is a propensity for measuring results to give a particular result. Of course, this might not seem like a particularly good explanation of the fact that the different methods were found to agree, but it does seem to be at least a possible alternative explanation. And if so, the question arises: why ought we to accept the hypothesis that the mountain has a physical height of 15, 873 feet in preference to it? If we say the explanation postulating an objective physical height is to be accepted because it provides us with the *better* explanation, then we are led back to our original question: Why ought the fact that some explanation is the best be a reason to believe it?

In summary, if what is in question is a realist (as opposed to say, operationalist) view of space, then appealing to the fact that two methods give us the same result seems misdirected. It is, to use a much-quoted passage from Wittgenstein, as if a man were to buy several copies of a newspaper to convince himself that what it said was true. The agreement between different methods is pretty evidently appropriate as a response to certain kinds of doubt, but it is not so clear that it is a well-directed response to all kinds of doubt. There is also, it seems, the danger that it implicitly appeals to some form of IBE if it says a particular hypothesis is to be preferred on the grounds that it offers the best explanation of the agreement between the independent methods.

The example given above refers to two ways of measuring the height of a mountain. But of course the same point could be made in any number of ways. The concept of valency in chemistry can serve as an example. The valency of chlorine, for example, could be found to be −1 from the way in combines with a range of other elements, or from electrolysis. The agreement between the different methods might furnish us with good reason for saying the valence of chlorine really is −1 rather than some other number. But these observations *by themselves* might leave a great deal of room for disagreement about the underlying theoretical account of what it is to have a valency of −1. Or again, the frequency of, say, yellow light might be measured by an interferometer, or means of the pattern formed in the two-slit experiment. The different methods might agree on the frequency, but a great deal of room might be left for competing views of exactly what it is that has that frequency.

Possibly the most commonly used argument for Scientific Realism from the agreement of different methods concerns Avogadro's number and the existence of molecules.[42] The fact that a wide range of ways of determining the value of

[42] Avogadro's Number of a given type of molecule is one "mole" of molecules of that type. Unhelpfully, however, a "mole" of a given type of molecule is standardly defended in textbooks as Avogadro's Number of those molecules. A more helpful way of explaining Avogadro's Number is as follows: one gramme of hydrogen is said to contain Avogadro's Number of hydrogen atoms. Similarly, 12 grammes of carbon-12 contain Avogadro's number of carbon atoms. If an element has atomic weight N. then N grammes of that element contain Avogadro's Number of atoms of that element. The value of Avogadro's Number is about 6.022×10^{23}.

Avogadro's Number all give the same value has been widely claimed to constitute good reason to believe atoms and molecules are real. Perhaps the most well-known scientific defence of the existence of atoms making reference to these methods is J. P. Perrin's *Atoms*. In considering the efficacy of this empirical work in establishing the existence of atoms, it is worth distinguishing two questions:

1. Does the empirical work of Perrin and others succeed in making a good case for the existence of atoms?
2. Is the method of agreement between independent methods by itself sufficient to establish a good case for the existence of atoms?

In this book it will be argued that the answer to the *first* question is "Yes, Perrin's work does constitute a good case for the existence of atoms." This is argued in some detail in Chap. 7. But here it will be argued that the answer to the *second* question is "No, the agreement between independent methods is, *by itself*, *not* enough to constitute a good argument for realism."

There are many ways of determining Avogadro's Number. And yet, all these different methods give – within the limits of experimental accuracy – the same value. All the methods lead us to say there are (approximately) 6.022×10^{23} molecules in one mole. This number is standardly denoted by "$\mathbf{N}_A$". On the face of it, the agreement between these methods seems to indicate there is *something* real being measured here. And, many have claimed that the things that are real are atoms and molecules.

We describe in detail some of the methods actually used to determine Avogadro's Number in Chap. 7. But here we will note some general problems with the argument.

If a wide range of independent methods agree on the value of $\mathbf{N_A}$ it certainly seems extremely plausible that, say, one gram of hydrogen has $\mathbf{N_A}$ "somethings" in it. But need these "somethings" be *atoms*? As we have noted, we might explain the agreement between different ways of determining the height of a mountain as due to the fact that the mountain did exist as a physical object with a certain height, or as due to the existence of certain propensities. The agreement between the two methods perhaps establishes the existence of a certain number of "somethings", but those somethings need not be feet considered as units of objectively existing real physical space. We can give another example. We may verify in a number of ways that a one-litre tin of paint always covers one hundred square metres. We may verify this by using a brush, or using a roller, or using a spray gun: all the methods result in the contents of a one-litre tin covering exactly one hundred square feet. But only in a rather stretched sense of the term does a one litre tin "contain" one hundred square feet; rather, it contains one hundred ten-cubic-centimetre volumes of paint, and these are sufficient to cover one hundred square feet with paint. The fact that different methods agree on some number N perhaps gives us good reason to say that there exist N entities of *some* sort, but it might leave it undetermined just what the nature of the entities may be. The agreement between the different methods shows that one gram of hydrogen contains $\mathbf{N_A}$ *somethings*, but this is perhaps not enough to show that these "somethings" must be atoms in the sense of small bits of matter.

It is worth examining this type of worry in a little more detail. Consider again the one litre tin of paint. It contains one hundred ten-cubic-centimetre volumes of paint. We would not ordinarily say that a ten cubic centimetre volume of paint *was* a square foot of paint. It need not even be the case that the volume of a square foot of paint is ten cubic centimetres: perhaps the physical operation of applying the paint to a surface increases or decreases the volume. Rather, what we can say is that ten cubic centimetres of paint in a can has a *disposition to produce* one square foot of paint, and that this disposition is independent of the method of paint application used.

It is clear that we can interpret the fact that there are a number of ways of determining Avogadro's number in a parallel way. The fact that different ways of determining the value of Avogadro's number produce the same result might be explained by, for example:

Each gram of hydrogen contains N_A hydrogen atoms.________(6)

But it could also be explained by something like:

Each gram of hydrogen has within it Avogadro's number of Xs, where each X has an operation-independent disposition to yield the same "unit" result in a determination of Avogadro's number.__________________(7)

Statement (7) is not quite the same as (6). To say something is an atom or molecule is, at least, to say that it is very small, discrete bit of matter. It is at least to say it has many of the properties of other bits of matter, like grains of sand or billiard balls, but, unlike those others, is too small to see. I think it is *ordinarily* a part of our concept of an atom that the discreteness of these tiny bits of matter is something that holds independently of the experimental tests we perform on them.[43] But (7) says much less than that. It merely tells us there is "something" that has a disposition to yield a particular experimental result, without specifying the nature of that "something". It might be tiny, discrete bits of matter, but it might be something else.

---

[43] I think this is part of the ordinary, lay person's conception of an atom. Of course, perhaps the "collapse of the wave packet" shows that in some sense the discreteness of a photon *is* somehow a by-product of the operations we perform on it, and does not exist independently of us.

The relation that the quantum theoretic "collapse of the wave-packet" due to the influence of the observer has to the issues presently under consideration seems to me to be very complex. However, I think it is fair to say that that the "core" content of "There exist atoms" or "There exist molecules" is something like: "There exist tiny bits of matter, with the properties that bits of matter generally have, and which are responsible for certain observable effects." Perhaps atoms and so on are subject to certain "weird" quantum theoretic effects. But then, so are *all* bits of matter, although the detectability of those effects at the level of things like tables and chairs may be much less. If all bits of matter are subject to these effects, then the assertion that there are bits of matter too small to see, and with the same properties as bits of matter we can see, remains unaffected.

It is worth noting that Pierre Duhem expressed doubts about atomic theory of just this sort. Duhem wrote:

> A chemical formula doesn't describe what really now subsists in the compound, but what can be found potentially, and can be taken out by the appropriate reactions.[44]

Here Duhem seems to be saying that if a chemical formula contains apparent reference to, for example, an atom of sulphur, then it ought not to be taken to be literally referring to an atom of sulphur considered as an entity, it rather ought to be seen merely as a reference to the potentialities or dispositions the compound has to respond in certain specifiable ways under the appropriate experimental conditions.

It might, perhaps, be claimed that (6), above, is a much *better explanation* of the agreement between the different methods than is (7). And, on the face of it, (7) does not seem to be a very good explanation. It "explains" why the different methods yield the same value by saying that an X is *disposed* to give the same "unit" result no matter which method is applied to it. On the face it, such an explanation of the agreement between the methods would seem to have about as much explanatory merit as Moliere's *dormative virtue*. (7) does not appear to be a very good explanation at all.

So, it seems quite clear that (6) is a *better* explanation of the agreement than (7). Hence, perhaps, we ought to *prefer* (6) to (7). But: does this mean that we are justified in accepting (6) as true? To say we are thereby justified in accepting (6) as true is, it seems clear, to rely on inference to the best explanation. And we have yet to see how we might be justified in doing that.

It is natural here to appeal to a notion of "contrastive confirmation".[45] The agreement of independent methods would seem to confirm there are 6.022×10^{23} "somethings" in a mole, *rather than* some other number of those "somethings", but it need not confirm that those somethings are tiny bits of matter, *rather than*, for example dispositions to produce certain observations.

In summary, when independent methods agree, for example, on the value of some quantity, a surprising event has occurred and some sort of an explanation is called for. But, there might be available a number of different explanations. One possible explanation, for example, might simply posit the existence of certain dispositions. And so, the question arises: What, if anything, justifies in accepting one particular explanation? If we say one particular explanation ought to be accepted because it is the best, we are simply back to our original question: "What justification do we have for inference to the best explanation?" Clearly, what we require is some sort of inference or argument that entitles us to say *more than* merely that there exists a disposition for certain results to be obtained. We want, for example, something that entitles us to assert that it is tiny bits of matter that are responsible for the

[44] See *Mixture and Chemical Combination and Related Essays* by Pierre Duhem, edited and translated with an Introduction by Paul Needham, *Boston Studies in the Philosophy of Science*, v 223, (Kluwer Academic Publishers, 2002), p. 92.

[45] For a discussion of the notion of contrastive confirmation see Jake Chandler "Contrastive Confirmation: Some Competing Accounts" *Synthese* (2013), v190, pp. 129–138.

fact that many different techniques yield the same value for Avogadro's Number. In subsequent chapters it is argued that Eddington inferences can play this role.

4.15 Structural Realism

In the previous section we saw that to explain the agreement between the different methods of determining Avogadro's number by postulating molecules seems to go rather beyond the evidence. It was not clear what, if anything, entitled us to accept the explanation postulating molecules rather than a number of other alternatives. Perhaps all we were entitled to make is the more epistemically modest claim that there exists *something* that behaves in a particular way. We are, perhaps, simply *not entitled* to make any more specific claims about the nature of the entities with which our theories deal.

One currently popular such "epistemically modest" theory is structural realism.[46] It is customary to distinguish between ontological structural realism and epistemic structural realism. Ontological structural realism says that all that actually exists is "structure". Epistemic structural realism says, roughly, that all we can know about is the structure: we cannot know about the nature of the entities participating in the structure.

This section does not offer any *general* argument against structural realism. Rather, two limited points will be made. First, it will be argued that the "epistemological modesty" underlying epistemic structuralism need not in all cases be required, second it is noted that the view advocated here is not confronted by a difficulty that is faced by ontological structural realism.

The idea that we can know about the "structure" of a particular domain is sometimes expressed as the idea that we can know the laws governing it.[47] So, for example, to adopt an epistemic structural realist view of a theory of light might be to say we can have good reason to believe light waves obey Maxwell's equations, while refraining from making any claim about the *nature* of light waves.

Epistemic structural realism, I think it is fair to say, makes the following two claims:

(i) We *cannot* have good reasons for claims about the nature of the entities dealt with by our theories.
(ii) We *can* have good reasons for claims about the laws obeyed by the entities dealt with by our theories.

[46] For a general survey, see James Ladyman "Structural Realism" in the *Stanford Encyclopedia of Philosophy*, ed. Edward N Zalta (2014) http://plato.stanford.edu/archives/spr2014/entries/structural-realism/

[47] I take it this is the position of John Worrall "Structural Realism: The Best of Both Worlds?" in *Dialectica* vol 43 (1989), pp. 99–124.

One of the central aims of this book is to argue that, in at least some cases, we can do better than this. More specifically, it is argued that in some cases we can have good reasons for claims about the nature of theoretical, unobservable entities. It is argued that Eddington-inferences give us good reasons for such claims. If this is granted, the epistemic modesty or scepticism that forms the motivation for (i) is undermined.

Ontological structural realism requires that it be possible to draw a distinction between claims that are *merely* about structure and other claims “over and above” those about mere structure. I think it is fair to say that it is controversial whether such a distinction can be satisfactorily drawn. But on the view advocated here, it is not required that it be possible to draw such a distinction. On our view, ontological commitments follow from epistemological claims. If an Eddington inference justifies us in saying that there are, for example, molecules considered as tiny bits of matter, then we have good reason for saying there are molecules in that sense, even if saying so involves making claims about “nature” over and above “structure”.

4.16 IBE Contrasted with the View Advocated Here: A Summary

In this chapter criticisms have been offered of IBE. It has been argued that the fact that some theory gives us the best explanation of some observations is not – at least if the theory makes claims about the existence and behaviour of unobservables – a good reason to accept the theory as true. But it might perhaps be protested that there is a sense in which the view adopted here does covertly use, or at least accept, a form of IBE. Let us consider again the discussion of induction given in Chap. 2. We there assumed that every crow we have observed in Geelong had been black, and it was argued that this gave us probabilistic reason to prefer the hypothesis that all crows are black. But now, saying that all crows are black surely furnishes us with an explanation of our observation. The hypothesis “All crows are black” *is* an explanation. But it has also been argued that it is better than the others since it is more probable. Therefore, it may be suggested, the view advocated here urges us to accept what is in fact the *best explanation* of our observations (“All crows are black”). So, is it not then the case that the view advocated here uses IBE after all?

Our aim in this book is to argue that we have good reason to accept the truth of Scientific Realism. The strategy adopted is to argue that scientific realist claims, or at least some of them, can be given a probabilistic justification. An initial statement of this probabilistic justification was given in Chap. 2 and it will be further developed in Chap. 5. The aim of the present chapter, however, has been to consider whether it might be possible to give a justification scientific realist claims *distinct from* the justification given in Chaps. 2 and 5. More specifically, the aim has been to consider whether realism might be justified via IBE, where what it is for an explanation to be “best” is for it to have some virtues distinct from being justifiable in the probabilistic way defended here. Such other virtues might include simplicity, unity, symmetry and so on. The aim of this chapter has been to argue that *these* other virtues are not "up to the job" of justifying realism. Of course, an explanation that can be given the strongest probabilistic justification of the sort advocated here might on those grounds be called "best", but this would not constitute a vindication of IBE in the sense with which we have been concerned in this chapter. It might also be protested that the probabilistic justification given in Chap. 2 is rather weak, in that it does not raise the probability of the preferred hypothesis by very much. But we have yet to see how IBE, in the sense with which we are concerned, raises the probability of claims about unobservables at all.

Chapter 5
On the Inference to Unobservables

The aim of this chapter is to explore and defend a route to realism about unobservable entities that does not use inference to the best explanation.

In Chap. 2, it was argued it is possible to give a justification of induction. It was there argued that a universal generalisation (such as "All crows are black") is *more likely than* other inferences that can be made from the premise "All observed crows are black". An *a priori* but defeasible probabilistic justification of induction was given. A main aim of this chapter is to argue that a similar kind of justification can be given for some inferences to unobservables. The justification is probabilistic and does not appeal to any principle of inference to the best explanation.

There is a desideratum that needs to be met by any satisfactory argument for realism:

> *Desideratum of Epistemic Sufficiency*: Our argument for realism must not merely furnish us with *some* reasons for realism; it must furnish us with sufficient reason to render realist claims worthy of rational belief.

One of the aims of the present chapter is to argue, in at least *some* cases, the inference to unobservables to be defended here does meet the above desideratum of epistemic sufficiency.

A version of this chapter has been given as a seminar in a number of places, including the University of Newcastle NSW, the University of Melbourne, Bristol University and the University of Athens. I am indebted to Joe Mintoff, Russell Blackford, Howard Sankey, Michel Ghins, James Ladyman and Stathis Psillos for helpful comments.

J. Wright, *An Epistemic Foundation for Scientific Realism*, Synthese Library 402, https://doi.org/10.1007/978-3-030-02218-1_5

5.1 Eddington's Fish Net

The scientist Arthur Eddington once asked us to imagine a hypothetical ichthyologist studying the size of fish.[1] He gathers fish in nets and studies the size of the fish that have been caught. These nets, we will suppose, have holes that let fish – if there are any – of less than two inches long through the holes. So, all the fish remaining caught in the nets will be at least two inches long. The ichthyologist, we will suppose, makes the observation that *there are no fish in the nets* less than two inches long. Ought the ichthyologist draw the conclusion *there are no fish in the sea* less than two inches long? From a common-sense point of view, of course, such a conclusion would not be justified. But Eddington himself argues against this common-sense conclusion. He defends this position by embracing a position resembling neo-Kantianism or, perhaps, what we might nowadays call a form of "internal realism".[2] According to the position Eddington embraces, it is only permissible to say "Xs exists" from within the conceptual framework of science. For Eddington, it is not permissible to say there might be entities out there not capable of being "caught" by the net of our language or concepts. So: Eddington says the ichthyologist ought to deny, or at least refrain from asserting, that there might be smaller fish in the sea.

Here we will not be concerned with Eddington's neo-Kantian or "internal realist" – if that is what it is – perspective.

Consider the hypothesis "There are no fish in the sea smaller than two inches". As we have noted, from a common-sense point of view we would not be justified in drawing this conclusion. Rather, I think common-sense tells us that, given these observations, we may actually have good reason to suppose *there were* smaller fish in the sea. More specifically, if the fish in the nets ranged in size from exactly two inches to rather larger, I think we would have good reason to say there *probably were* fish in the sea smaller than two inches in length. We can, moreover, give a justification for this conclusion. The reasoning proceeds as follows:

> Suppose there were no fish in the sea smaller than two inches. Then, a highly improbable event would have occurred: the blindly chosen size of the holes in the net would have

[1] See Arthur Eddington *The Philosophy of Physical Science*, Tarner Lectures, (Cambridge University Press, 1939) pp. 16–18.

[2] See Eddington, *loc cit.* Eddington writes "Anything uncatchable by my net is *ipso facto* outside the scope of ichthyological knowledge. In short, what my net can't catch isn't fish."

It is perhaps worth stressing that Eddington himself *did* believe in things too small to see with the unaided senses. For example, he believed in the existence of protons and neutrons, and much of his cosmological work concerned how the numbers of these particles might determine the nature of the universe. His story of the fish net is not intended to cast in to doubt the existence of things not detectable by our senses. In his story, the fish net does not correspond to our eyes or ears, but to our *language or concepts*. Eddington is (perhaps) merely pointing out that any claim we make must always be made from our scheme of concepts. The position he is opposing is neither instrumentalism nor what we might nowadays call constructive empiricism. What he is opposing is possibly rather closer to what is sometimes called "Metaphysical Realism", or the idea – rejected by Nelson Goodman, for example – of a real world underlying all our versions of it.

> happened to have coincided with the size of the smallest fish in the sea. But since this is unlikely, we have good reason to believe the size of the holes in the fish net is not the same as the size of the smallest fish in the sea. We know the size of the smallest fish in the sea is not larger than two inches, since our net contains fish two inches and longer. Therefore, we may conclude, there are probably fish in the sea smaller than two inches.

It is useful to have a name for this inference. We will call it the inference from Eddington's Catch to the Existence of Smaller Fish, or the EC-ESF inference. Intuitively, it seems to be an inference with some merit.

At least on the face of it, the prospects for the EC-ESF inference being able to meet the requirement of epistemic sufficiency seem quite good. That there are fish in the sea of two inches *or more* certainly meets this condition. Assertions like "This is a fish", "This is a table", "This is a coffee-cup" are, under the right conditions, perhaps the paradigms of propositions that are worthy of rational belief. Even if we do not uncontroversially possess a full understanding of that in virtue of which they are rational, it seems plausible to say philosophy ought to accept that these are things which, under the right circumstances, it is rational to believe. As we have noted in earlier chapters, there is a tradition in philosophy, often associated with G. E. Moore, of seeing claims like these as more firmly based than the premises of sceptical arguments.[3] But now, on the face of it, we seem to be presented with an argument of the following form:

Premise (1): Fish of two inches or more exist.
Premise (2): If fish of two inches or more exist, then probably fish of less than two inches exist.

Therefore, probably: Fish of less than two inches exist.

The first premise of the argument is, it seems, a paradigm of the type of belief worthy of rational acceptance. The argument, if sound, shows that there is a *probabilistic* link between this paradigmatically rational belief and a belief about a class of entities that is, in a sense, unobservable. The argument, if sound, would seem to show that the difference in rational acceptability between "There are fish we can detect" and "There are fish we cannot detect" is only one of degree and not fundamental kind.[4]

---

[3] What is widely regarded as the classic statement of this position is given in G. E. Moore's papers "A Defence of Common-Sense" in Contemporary British Philosophy edited by J. Muirhead (1925) and "Proof of an External World" in Moore's *Philosophical Papers* (Routledge, 1959) Chapter 7, pp. 126–148.

[4] The approach advocated here has some points of similarity, as well as some points of difference, to the approach sketched by M. Ghins towards the end of his "Putnam's No-Miracle: A Critique" in Lyons and Clarke (2002), pp. 121–138. Ghins accepts that existence-statements such as "This table exists" are rational. But, for Ghins, it is a philosophical task to understand or explicate just what it is that makes such existence statements rational. He suggests that the way to establish scientific realism is to first develop an understanding of that in virtue of which statements like "Tables exist" is rational, and then proceed to argue that a justification of the same sort can be given for "Electrons exist".

It should be noted that there are many beliefs that are surely worthy of rational acceptance but which are such that they have *not been observed* to be true at the time of utterance. One such belief might be "The Sun will rise tomorrow". Of course, such a belief is *verifiable*: all we need to do is wait 24 hours to verify it. But, at the time of utterance, it has not been verified. And yet, it is surely a rational belief *at the time of utterance*. The question therefore arises: In virtue of what is it a rational belief at the time of utterance? A natural reply is: Because there is a probabilistic link between this assertion and other events that have been observed. But, at least *prima facie*, there also seems to be a probabilistic link between the existence of detectable fish and fish too small to be detected. If a link of this sort is sufficient to make a proposition about the rising of the sun worthy of rational belief, it is hard to see why it would not be sufficient to do so in the case of the fish too small to catch.

5.2 Eddington's Inference and Induction

The inference from the caught fish to the fish too small to catch appears to be a probabilistic inference. However, *prima facie*, at least, it does not look very much an (enumerative) inductive inference. It does not, for example, look very much like the inference from "All observed crows are black" to "All crows are black".

It is useful to get clear on the ways in which the EC-ESF inference and a typical enumerative inductive inference *seem* dissimilar. There are perhaps two main ways:

(i) When we make an inference from "All observed crows are black" to "All crows are black" we conclude that unobserved crows are like observed crows (They're black!). But when we make the EC-ESF inference we seem to be doing something very different. We make an inference from "All observed fish are two inches or longer" to the conclusion "There are unobserved fish *less than* two inches long." On the face of it, here we seem to be saying unobserved cases are *unlike* observed cases.

(ii) There is another respect in which the two inferences seem to be dissimilar. The inductive inference takes us from "All observed crows are black" to the *universal generalisation* "All crows are black". Universal generalisations are generally thought to have no "existential import": to say "All crows are black" does not imply there are any crows. But the EC-ESF inference takes us from the observation of fish two inches and more in length, to the conclusion that there (probably) *exist* fish less than two inches long. The conclusion of the EC-ESF inference *does* tell us something exists.

Given (i) and (ii), it would seem that the EC-ESF inference and induction were rather different. But in what follows it will be argued that they are in fact more similar than they seem. More specifically, it will be argued that while (ii), above does pick out a genuine difference between the two inferences, (i) does not. Most importantly, however, it will be argued that if we accept the justification of induction given earlier, we ought rationally also to accept the EC-ESF inference.

We have just given a brief sketch of how the EC-ESF inference might naturally be justified. This can be compared with the argument for induction given in Chap. 2. To briefly recapitulate that argument: Suppose we have observed all of the crows in one particular region – Geelong, let us say – and they have all been black. Does this give us reason to suppose all crows (in other regions of space and time) are also black? Suppose it were not the case that crows elsewhere were black. More specifically, suppose it were the case that while all crows in Geelong were black, all the crows elsewhere were green. If so, a highly improbable event would have occurred: the blindly chosen location of our observations would have happened to have coincided with the region of black crows in a sea of non-black crows. Since such an event is highly improbable, we have reason to say it has not occurred. That is, it is probably not the case that Geelong is an island of black crows in a sea of non-black crows. But of course, to deny Geelong is an island of black crows in a sea of non-black crows is not to say all crows everywhere are black. There are many possibilities: Perhaps only a fraction of the crows outside Geelong are black, or perhaps crows prior to our observations were all green and afterwards they were all black, or perhaps crows everywhere in the world are black except for some small, restricted region. (Perhaps crows in Stockholm are green and everywhere else they are black.) All of these are, of course, possibilities, but it was argued in Chap. 2 that there is a clear sense in which all of them are less likely than "All crows are black". For all of them, the blindly chosen location of our observations (Geelong) would have happened to have coincided with a location in which all the crows are black. Unless all crows, at all points in space and time, are black, the probability of this happening is less than one. The only hypothesis that maximises the probability of our observations having taken place is the hypothesis "All crows are black".

In Chap. 2 the thesis was defended that the above argument for induction is a good one. Here it will be argued that, given the above argument for induction, we also ought to accept the EC-ESF inference to the existence of the smaller fish.

5.3 Eddington Inferences and Induction: Similarities and Differences

Let us begin by noting the similarity of the logical structure of the two arguments. Both arguments begin by asking us to assume the negation of that which is to be proved. In the case of induction, we begin by assuming it is not the case that all crows are black, and with EC-ESF we assume it is not the case that there are fish smaller than two inches in the sea. In both cases these assumptions lead us to say that the blindly chosen locations of our observations must have been improbable (or in the case of induction, less than maximally probable). In both cases, this is seen as a reason for preferring the original hypotheses (All crows *are* black, There *are* smaller fish) to their negations. The logic of the two arguments are the same: if we accept the one in support of induction – and the second chapter it was argued we ought to do this – then it seems we also ought to accept the EC-ESF inference.

However, the similarity of the two arguments perhaps makes their differences all the more puzzling. As we have already noted, the two arguments lead us to – apparently – very different types of conclusion. The conclusion of the EC-ESF inference leads to an existence-claim, the conclusion of the inductive inference does not. The EC-ESF inference leads us to say that unobserved cases are unlike observed cases, but the inductive inference tells us the unobserved will be like the observed. In what follows it will be argued that the first of the two differences (to do with existence) is real but does not undermine the EC-ESF inference. The second difference is actually illusory. In both cases, the cogency of the EC-ESF inference remains intact.

We will begin by considering the second apparent difference: that, in contrast to induction, the EC-ESF seems to lead us to say the unobserved is *unlike* the observed. It will be argued that this is not so: properly understood, the EC-ESF and related inferences *do* lead us to say the unobserved will be (relevantly) like the observed.

We can begin by imagining a modified version of Eddington's example. Suppose we have a fish trap that has an opening that only allows fish of a certain size (if they exist) to enter. We could, for example, set the size of the opening so that only fish of *exactly four inches* in size can enter. But, let us further suppose, we can alter the size of the opening in the trap. We could, for example, set the size of the opening so that only fish of, say, exactly five, or exactly six, inches get in. Also, let suppose, we can move the trap around. We can move it to different geographical regions of the seas and oceans, different depths and so on. More generally, we can say there are two types of "setting" for the trap: We can set the size of the opening of the trap, and we can also set its location in the sea. Both types of setting are chosen "blindly" – that is, in ignorance of the existence and properties of any fish we might catch.

Now, let us suppose we *blindly* set the trap, in the following senses: we place it in region R of the sea, and we set the size of the opening to exactly four inches. Both settings are blindly chosen in the sense that we do not know if there are any fish in R, neither do we know if there are any fish of exactly four inches. After a while, we open the trap and find there are fish in it. Of course, the fish will be four inches long. They will also be from region R. Let us also assume one more thing about the fish: they all have scales. What conclusions may we draw from all this? One obvious conclusion we may draw is that fish *exist*. We can also say: fish exist in region R, and fish exist four inches long. But there are other inferences we can make.

One way we can reason from our observations is as follows: it is highly unlikely that it should be merely due to chance that the blindly chosen size of the opening in the trap should be exactly the same size as all the fish that exist; therefore, there are probably fish of sizes other than four inches. Does this involve saying the unobserved is "unlike" the observed? Obviously, in one way it does: it is saying there are fish that, compared to fish we have observed, are unlike them *with respect to size*. But in another respect, it is saying that future observations will be *like* our present observations: Even if we change one of the settings on our fish trap (specifically, the size of the opening), the result will be the same: we will still get fish in our trap.

We get precisely parallel results if we change the other "setting" on the fish trap: its location in space. We place the trap in region R and get fish. Since it seems unlikely that the blindly chosen location R should be the only region with fish, we are led to predict that if the setting of the trap – in this case its location – had been changed, we would still have got fish. Are we here predicting that future observations will be "like" or "unlike" currently obtained observations? And again, the answer is: in some ways "like", in other ways "unlike". In one sense the results are the same: we get fish. But in another sense, they are unlike: the fish come from a different location in space.

In summary, whether we change the "setting" of the size of the opening, or the location in space, the results of the new observations have a similar relation to the earlier observations. They are similar in that they are all observations of fish. But they are also different: in the first case, the difference is one of size, in the other case it is one of location. In both cases, the difference is, of course, due to the nature of the difference of the settings.

There is another point to note here. Suppose we put our trap in region R, with size of the opening set at four inches, and we get fish in the trap. Then there are at least two inferences we can make:

(1) There are probably fish of lengths other than four inches in the sea.
(2) There are probably fish in the sea in regions other than R.

These two inferences can be obtained from the same body of data. What justifies the first inference is the fact that the setting on the fish trap of an opening size of four inches is blindly chosen. What justifies the second inference is the fact that the setting on the trap of location R is blindly chosen. The inferences we can make from a body of data are not merely determined by the data themselves, but by the respects in which the settings on the trap were blindly chosen.

Of course, we have yet to make what we would typically call an inductive inference from the data. The inference we have (so far) made from the fact that we have blindly chosen location R for our trap is that is that *there are* probably fish in the sea in locations other than R. But in an inductive inference we move from "All observed As are Bs" to "All As are Bs". The conclusion of an inductive inference is "universally quantified", while the conclusion of the other inference is "existentially quantified". But we can make some inductive inferences from the contents of the trap. As we noted before, perhaps the fish we got in our trap had scales. So, we could make the inference:

<u>All observed fish have scales</u>
All fish have scales

A way of justifying this inference has been defended. (Briefly: all fish from blindly chosen region R have scales, it is highly unlikely that that our blindly chosen location should have happened to have coincided with the one and only location of scaled fish. So, the hypothesis to be preferred is that all fish everywhere have scales.) But now we seem to be confronted with a puzzle: If the inference from "All observed fish have scales" to "All fish have scales" is a good one, why isn't the following inference also good:

All observed fish come from region R
All fish come from region R.

However, the first but not the second inference can be given a justification of the sort we have been using. To justify the second inference, we would need the premise: "It is highly unlikely that the blindly chosen location of our observations (location R) should have coincided with the observed property of the fish (coming from location R). But, obviously, this is not "highly unlikely": it is in fact trivially and necessarily true. So, a justification of the second inference, of the sort we have been using, cannot be given.

The thesis has been defended that the inference from the contents of the fish trap to the conclusion "All fish in the sea have scales" and "There are fish in the sea of lengths other than four inches" can be justified by arguments using the same "underlying logic". It has also been argued that some, but not all, of the differences between the two conclusions are illusory. Of course, one difference is not illusory: the conclusion of the EC-ESF inference has "existential import", the conclusion of the inductive inference does not. But it is natural to account for this difference in terms of the starting points of the two inferences, not as a difference between the "logics" of the subsequent justifications. We can make this clearer:

We put our trap in the sea and get some fish. There are fish in our trap, they all have scales, they are all four inches long and they are all from region R. Call this our data. One set of conclusions we can immediately draw from this data make reference to *existence*: "There are fish", "There are fish four inches long", "There are fish in region R" and so on. The relation between our data and these conclusions is deductive, not ampliative. But there are other conclusions we can draw: "All fish in the net are four inches long", "All the fish in the net have scales", "All the fish in the net come from region R". The relation between the data and these descriptions is also non-ampliative. But from these two different starting points the reasoning process is the same. In all cases the reasoning proceeds as follows:

> The question is asked: "Is P true?" It is noted that if P were true an improbable event would have occurred: the setting on the fish trap would have coincided with the distribution of some property in the world or some feature of the world. The inference is made that since this is improbable, it is probable that not-P is true (or at least that not-P has a higher probability than alternatives.) Depending on the initial deductive step from the data, reasoning leads us either to some conclusion such as: "There are fish of lengths other than four inches" or an apparently quite different conclusion such as "All fish have scales".

Let us now summarise the results of this section. We have here been concerned to defend the inference from "Our trap is set to four inches, and there are fish in the trap" to the conclusion "There are probably fish of lengths other than four inches in the sea". Specifically, we have been concerned to defend the thesis that this type of inference is as good as an inductive inference and can be justified in the same way. So: if were prepared to accept inductive inferences, so ought we to be prepared to accept the EC-ESF inference.

5.4 Eddington Inferences More Firmly Based than Induction

It is worth noting that there are respects in which the EC-ESF inference would actually seem to be on an epistemologically firmer footing than standard inductive inferences. Standard inductive inferences are of the form: “All observed As are Bs”, so “*All* As, everywhere and everywhen, are Bs”. In the example under discussion, an inductive inference might say: “All observed fish have scales, therefore: All fish have scales”. The conclusion is a universal generalisation. But the conclusion of the EC-ESF inference is an existence statement: there is exist fish smaller than four inches. The *a priori probability* of the universal generalisation is, it would seem, very low. In an infinite universe it would presumably be zero. The *a priori* probability of the existence statement is, it intuitively seems, much higher. In an infinite universe it would presumably be one. And yet, the two inferences, it has been argued, start from the same empirical evidence and proceed via the same inferential steps. The inductive inferences leads to a conclusion with a lower *a priori* probability, the EC-ESF inference to a conclusion with a higher *a priori* probability. This surely shows that the EC-ESF inference is on a stronger epistemological footing than the standard inductive inference. If we accept induction we surely also ought to accept the EC-ESF inference.

It is appropriate at this point to return to a matter raised in Sect. 2 of Chap. 2. We there noted that, although inductive support confers an increased probability on generalisations, the level of probability conferred may be low. However, it was also claimed that this need not threaten the approach to justifying Scientific Realism advocated here, since (it was claimed) Eddington inferences are stronger than inductive inferences. And we have just seen that this is so. Eddington inferences, it has been argued, confer a higher probability than do inductive inferences. Moreover, it will be argued, in at least some cases Eddington inferences are strong enough to yield what we are inclined to call “knowledge”.

There are some inductive inferences that yield conclusions we are strongly inclined to say we “know”, and which are also, I think, plausibly amongst the paradigms of things worthy of rational belief. An example might be “The sun will rise tomorrow”. It will be argued that the conclusion of the EC-ESF inference gives us a belief of a comparable degree of rational support.

Let us consider in a little more detail the nature of the rational support for “The Sun will rise tomorrow”. We have observed the Sun rising on many occasions. Will it rise tomorrow? Suppose it will not. Then a highly improbable event would have occurred: our current point in time would happen to have coincided with that point in time at which a non-rising of the sun was to take place amidst a sea of risings. It might perhaps be suggested that it need not be the case that there is *just one* rising amidst a sea of risings. Perhaps from now on the sun will not rise. But then, a highly improbable event would still have taken place: our current point in time would have happened to have coincided with the point in time at which risings changed to non-risings. It is, I think, for reasons like this that we say “The sun will rise tomorrow” is worthy of rational belief. But the reasoning that supports “There are fish in the sea less than four inches" would appear to be *at least* as strong as that supporting our belief that the sun will rise tomorrow. In fact, our grounds for saying "There are (some) fish in the sea less than four inches" would appear to be of a comparable degree of strength to that which would support "There will be a rising of the sun at some point in the future". Since this latter belief is logically entailed by "The Sun will rise tomorrow", it would seem to follow that the case for "There are fish in the sea less than four inches" must be at least as good as that for "The Sun will rise tomorrow". And since the latter is clearly worthy of rational belief, it follows that so must "There are fish too small to be detected by our trap" be worthy of rational belief. That is, there are at least some statements about unobservables that are worthy of rational belief. The approach advocated here can, in at least some cases, meet the requirement of epistemic sufficiency.

5.5 Eddington Inferences and Unobservable Entities

We may now draw some more general conclusions from the preceding discussion. First, let us introduce the more general notion of an "Eddington inference": named after the man who denied that such inferences are rationally permissible. We perform an "Eddington inference" if we begin by asking: "Is it the case that the only entities that exist, or have property P, are the ones we can, in some sense, observe?" We then argue that it is highly unlikely that this should be the case since, if it were, then the blindly chosen restriction on what we can observe would (improbably) have coincided with what exists, or with what has P. So: we conclude that things exist, or have P, that we cannot observe.

This type of argument can, of course, be applied rather more widely than our example of the fish net. It can be naturally generalised or extended to apply to a wide range of cases. Here is one such extension:

> Are there entities that cannot be detected with our unaided senses? In particular, are there entities too small to see? If not, a highly improbable fluke would have occurred. The size of the entities detectable on the blindly chosen "setting" of our observational apparatus (our actual senses) would have happened to have coincided with the smallest entities there are. Since this is unlikely, we may conclude it is (probably) not the case. Therefore, there (probably) do exist entities too small for us to see. (Note: the logic of the argument remains unaffected if we replace "too small to see with our unaided senses" with "too small to see with our most powerful detecting apparatus".)

It may be objected that the above line of argument only supports a fairly weak and unexciting form of realism. It does not say any *specific entities*, of the kind a realist is likely to be interested in, exist. (For example, it does not say *electrons* exist.) All it does is say that *some* unobservable entities, with *some* properties, exist.

However, even if the form of realism so far defended is fairly "weak and unexciting", it is a form of realism that can be justified in the same way that inductive inferences are justified. It is a form of realism that does not rely on IBE, but rather relies on something like induction. It seems therefore to provide us with a firmer route to at least some unobservables than does IBE.[5] The question naturally arises, though: Can we use Eddington inferences to establish more interesting cases of realism? It will be argued in a later section that we can, but first we must examine more closely some aspects of Eddington inferences.

5.6 Restricted and Unrestricted Eddington Inferences

Suppose we have found our net containing fish of a range of sizes down to four inches, this being the size of the holes in our net. Then we could draw the simple conclusion that there are fish in the sea less than four inches. No restriction is placed on the size of the uncaught fish, other than that they are smaller than the holes in our net. Call an inference of this sort an *unrestricted* Eddington inference. But there are other inferences we could make. We could draw the conclusion there are fish in the sea of less than three inches, or less than two inches, or less than one inch. These, we will say, are all *restricted* Eddington inferences.

The question arises: Are both restricted and unrestricted Eddington inferences good? Are some better than others?

It seems intuitively clear that unrestricted Eddington inferences are stronger than the restricted inferences. For example, Inference A:

All the fish caught in our net with holes four inches across are four inches or longer.

Therefore: There exist in the sea fish smaller than four inches.

Is clearly stronger than Inference B:

[5] In this book the position is adopted that induction and Eddington inferences are on firmer ground than inference to the best explanation. And the reason for this is because it was argued that both induction and Eddington inferences can be justified, whereas in Chap. 5 it was argued that we are not as yet in possession of a justification of IBE. However, in an influential paper, Gilbert Harman argued that enumerative induction ought to be seen as a special case of inference to the best explanation. (See G. Harman "The Inference to the Best Explanation" *The Philosophical Review*, vol 74, (1965), pp. 88–95.) On the face of it, if Harman's claim is correct, it would appear to be potentially very damaging to the position adopted here. However, it seems to me that the thesis Harman is arguing for is perhaps consistent with, and not a threat to, the position advocated here. When Harman suggests that enumerative induction is *best seen as* a special case of IBE, what he is suggesting is that when a typical speaker (of English, say) offers what looks like an enumerative inductive argument, what the speaker is *actually*, perhaps implicitly, doing is using IBE. What Harman is doing, that is, is offering something like a thesis about what is implicitly "going on in the head" of the typical speaker when that speaker gives an argument. Harman is not, as far as I can see, offering any kind of thesis about rational justification. But in this book no claim is made about what is "going on in the head" of a typical speaker when they argue. Our concerns are with rational justification. So, as far as I can see, the claims made by Harman and the claims made in this book are distinct, and compatible.

All the fish caught in our net with holes four inches across are four inches or longer.

Therefore: There exist in the sea fish smaller than three inches.

It is clear why Inference A is stronger than B. As we have already noted, if the conclusion of Inference A were false, a highly unlikely event would have occurred: the size of the smallest fish in the sea would have happened to have coincided *exactly* with the size of the blindly chosen holes in our net. Since this seems to be very unlikely, Inference A is very strong.

Inference B is not as strong as Inference B, but it still does have some strength. Suppose the conclusion of Inference B were false. Then a somewhat improbable event would have occurred: The blindly chosen size of the holes in our net was close to (in fact, an inch or less) away from the size of the smallest fish in the sea. Given that, for all we *a priori* know, fish could come in a range of sizes, it seems somewhat unlikely that the smallest size of them should happen to be so close to the size of the holes in our net.[6] So, Inference B still has some probabilistic force, although rather less than Inference A.

[6] It might be suggested that the view of probability given here is inconsistent with the view given earlier. It has here been suggested that it is *a priori* unlikely that the size of the smallest fish in the sea should coincide with the holes in our net. From this it follows that it is *a priori* more likely that should be fish smaller than the holes than that there should be no such fish. But it might be felt this is inconsistent with the position adopted in Chap. 2, where it is stated that from an *a priori* point of view it is not more likely that our apparatus should display a particular numerical result (say, the needle pointing to "7") than it is that the apparatus should turn in to a bowl of petunias. However, it will be argued that there is in fact no inconsistency here.

First let us look more closely at the claim it is *a priori* unlikely that the size of the smallest fish in the sea should coincide with the size of the holes on our net. Suppose the holes in our net are 2 inches across. Then, given that our net contains fish of many sizes greater than 2 inches, there are many lengths the smallest fish might have. *One* of these possibilities is:

(1) The smallest fish are 2 inches
(2) But the following are also possibilities:
(3) The smallest fish are 1.9 inches
(4) The smallest fish are 1.8 inches
(5) The smallest fish are 1.7 inches

And so on.State of affairs (1) – that the smallest fish should be exactly 2 inches – is of course *only* one possibility amongst many. On these grounds it is therefore *a priori* less likely that this *particular* state of affairs given in (1) should obtain rather than that the disjunction "(2) or (3) or (4) and so on" should turn out to be true.

Let us now consider the *a priori* probability of the outcome of some experiment. Amongst the possible outcomes are:

(a) The needle points to "7".
(b) The device turns in to a bowl of petunias.

Of course, given our (empirically obtained) knowledge of how the world works, (a) would surely be more likely than (b). But such an assessment of likelihood is, of course, *a posteriori.* From a purely *a priori* point of view, there would seem to be no evident reason to regard (a) as more likely than (b). Crucially, the comparison between (a) and (b) is a comparison between two

Do Eddington inferences give us reason to believe there exist fish down to any arbitrarily small size? Do we, for example, have reason to believe there are fish of a millionth of an inch in length? It will be argued that we do, and the reasoning parallels the argument for induction given in Chap. 2.

Suppose we already know there are, say, fish a foot long in the sea, but do not know how small they get. The size of the smallest fish in the sea lies somewhere between one foot and zero, but we do not know where. We will say the *potential range* of the size of the smallest fish goes from one foot to zero. We choose the size of the holes in our net to be somewhere within this range, but the precise location is chosen blindly. Suppose we choose the size of the holes to be four inches. And: we get fish in our net, of a range of sizes, down to and including the size of the holes in our net. It is obviously highly improbable that the blindly chosen size of the holes in our net should have happened to have coincided *exactly* with the size of the smallest fish. So, the inference to the conclusion that there are smaller fish seems very strong. But it is also unlikely that the blindly chosen size of the holes in the net should happen to have been very close to the size of the smallest fish. It is, more precisely, unlikely that the size of the smallest fish should have happened to be within δ, for some small δ, of the size of the holes in our net. If it should happen that the size of the holes in our net was within δ of the smallest fish, an unlikely event would have occurred: the blindly chosen size of our holes would have happened to have fallen with a particular range. Of course, as δ becomes larger, the chances that the blindly chosen size of the holes should have been δ or less away from the size of the smallest fish becomes greater. But, if the size of the smallest fish is greater than zero, the probability that the blindly chosen size of the holes in our net would have been within this range is less than one. The probability that the blindly chosen size of the holes in our net should have fallen within this range is at the maximum value of one only if the size of the smallest fish is zero. Therefore, the hypothesis that the size of the smallest fish is zero, or at least infinitesimally close to zero, is to be preferred to the hypothesis that they are any other specific size.[7]

These considerations bring out how close is the relationship between enumerative induction and what we are here calling Eddington inferences. We could, in fact, regard Eddington inferences as the "existential counterpart" of induction. In both types of case, the justification of an inference relies on the fact that the location of our observations is "blindly chosen". In both cases, the justification proceeds by noting that if the conclusion of the inference were false and the location of our observations blindly chosen, an improbable event would have occurred. In the case of induction, there are a number of conclusions that might be drawn from the data:

---

specific outcomes or states of affairs. In this respect, it is unlike the other case just considered which involves comparison between a specific state of affairs and a *disjunction* of states. So, the reason for holding that it is *a priori* unlikely that the smallest fish in the sea should be exactly two inches does not hold in the "needle pointing to "7"" versus "bowl of petunias case".

[7] Of course, there will no doubt be other considerations (to do with physiology etc.) that will tell us there cannot be fish below a certain size. But here we are only concerned with the relation that exists between *one specific item of data* (the fact that we have found fish of four inches in our trap) and hypotheses about the existence of smaller fish.

that the fish in the vicinity of our trap have scales. These conclusions are: that the fish in this immediate vicinity have scales, that the fish within a hundred miles of the trap have scales, that the fish in this particular ocean have scales, and so on. But the most general hypothesis we can draw from our data is that all fish everywhere and everywhen have scales. And this conclusion is to be preferred to any other comparably specific hypothesis on the grounds that, if it were not so, a less than maximally probable event would have occurred with respect to the blindly chosen setting on the location of our trap. In the case of an Eddington inference, there are also a number of conclusions that can be drawn: that there exist fish of just less than four inches, that there exist fish ranging in size from three inches to four inches, that there exist fish ranging in size from two inches to four inches, and so on. The strongest hypothesis we can draw from the data is that there exist fish in all sizes down (infinitesimally close) to zero. And this hypothesis, too, is to be preferred on the grounds that, unless it were true, a less than maximally probable event would have occurred with respect to the blindly chosen setting on the holes of our trap.

5.7 Eddington Inferences and Partitioning

So far in this chapter the notion of "Eddington inferences" has been introduced and it has been argued that they can be given a justification like that given for induction. But, of course, our larger aim in this book is to argue that Eddington inferences help us to develop a theory of what counts as good reasons for scientific realism. More specifically, our aim is to develop a theory of what it is to have to have good reason to believe in the existence of the unobservable entities postulated by scientific theories.

But is it even possible to have good reason to believe in the existence of unobservable entities? The notion of "partitioning" might seem to suggest that we cannot.

Not all parts of a theory that explain some data receive evidential support from the data. Suppose, for example, a theft has been committed. The thief, let us suppose, gained entry by bending some iron bars across a window. As far as we can tell, the thief bent the bars with his bare hands. Then we may, perhaps, explain the observed facts of the robbery with the following hypothesis:

H1: The robbery was committed by a man with strong arms and who was wearing an orange hat.

H1, we will assume, does provide a possible explanation of the data, but not all parts of H1 would seem to receive evidence support from that data. The part of H1 that refers to the strong arms of the thief would appear to receive such support, but the part which asserts he was wearing an orange hat would not.

The point can be generalised. Let H be a hypothesis that explains some body of data D. Let C(H) be the set of all statements entailed by H. Then, even though H explains D, it need not be the case that every member of C(H) receives evidential support from D. Perhaps some do while others do not. And so, the question arises: which parts of H receive support from D, and which parts do not?

This question, of course, has relevance for Scientific Realism. Suppose H postulates unobservable entities. Are the claims that H makes about unobservable entities supported by D? There is an argument that suggests that they might not be so supported. Let us divide H in to its observable component H(O) and its unobservable, theoretical component H(T). Plausibly, all the observations that can be explained by H can also be explained by H(O). And so, the question arises: are we entitled to assert that both H(T) and H(O) receive support from the observational evidence? And if we are so entitled, Why?

It will be argued that the notion of an Eddington inference provides a promising approach for dealing with these difficulties. We can begin by noting that the evidential relation between the premises of an Eddington inference and its conclusion seems much "tighter" or "more directed" than that between the data *explained* by a hypothesis and the hypothesis itself. Consider again our example of an Eddington inference:

Premise: The fish caught in our net with holes four inches across exhibit a range of sizes, from four inches upwards.

Conclusion: There are probably fish in the sea of less than four inches.

The conclusion of the inference does not attribute to the smaller entities any property not possessed by the entities that we have observed. The only properties it attributes to the smaller entities are the properties associated with being a fish. And to attribute either more or fewer properties to the smaller entities would be to assert something not supported by the Eddington inference. Suppose that all the fish in our nets have scales. Then, an Eddington inference leads us to say that (probably) there exist smaller fish that have scales. To say that there are no smaller fish with scales would be to say a highly improbable event had occurred: it would be to say that the division between fish with scales and those without happened to have coincided with the blindly chosen size of the holes in our net. Eddington inferences lead us to say if the objects capable of being observed by us have properties $P_1, \ldots, P_n$ then there (probably) exist entities not capable of being observed by us that also have those very same properties.

These considerations bring out how the evidence supplied by Eddington inferences is more focussed and directed than that supplied by IBE. Each part of the conclusion of an Eddington inference is supported by evidence. An Eddington inference leads us to attribute *P* to unobservable entities only if we have observationally verified that there are some observable entities that have that same property *P*. If a conclusion is derived by an Eddington inference, no parts of the conclusion are, it seems, left unsupported by the evidence. The type of evidence supplied by Eddington inferences is not susceptible to the partitioning problem.

5.8 Eddington Inferences and the Paradoxes of Induction and Confirmation

It might be thought that the conclusions of the above section are a little too swift. Surely, it might be suspected, there may be *some* circumstances in which the use of Eddington inferences gives us conclusions parts (at least) of which are unsupported by the evidence. And the close relationship between Eddington inferences and induction might give us further grounds for this suspicion. After all, it is well known that there are cases in which inductive inference leads us to conclusions that are not, or would not seem to be, supported by the empirical data from which they are, putatively, inferred. So: might there not be situations in which Eddington inferences similarly lead us to conclusions not supported by the data?

Induction gets in to trouble with "gruesome" predicates.[8] The inductive move from "All observed emeralds are grue" to "All emeralds are grue" is plainly unsatisfactory, so it might be thought Eddington inferences get in to a related difficulty.

But actually, Eddington inferences do not get in to trouble here. Suppose the fish we have observed have green (and hence grue) scales. Then, the following Eddington inference can be made:

All fish caught in our net have grue scales.

Therefore: there exist fish too small to be caught that have grue scales.

Considered as an inference to a claim about what fish exist *now*, the above Eddington-inference is surely unexceptionable: it is clearly reasonable to believe there now exist smaller fish with green scales; therefore, it is reasonable to say there now exist smaller fish with grue scales. Of course, these smaller fish will presumably cease to be grue after D-day, but that does not refute the fact that *there are now* smaller fish with grue scales, and hence that the conclusion of the Eddington inference is true.[9]

There are other cases in which an inductive inference does not seem to support its conclusion, but where a corresponding problem does not arise for Eddington inferences. For example, the following inductive inference is presumably bad:[10]

---

[8] See Nelson Goodman *Fact*, *Fiction and Forecast* (Harvard University Press, 1954) p. 74.

[9] There are some Eddington inferences using "grue" that produce paradoxical results. The following is an example:

<u>All fish observed before D-day have been observed to be grue.</u>
Therefore, there exist fish after D-day that are grue.

This inference in effect takes us from an observation that all fish observed before D-day have been green, to the conclusion that there exist fish after D-day that are blue, and so the premise here certainly does not seem to support the conclusion. As far as the present author can see, this paradoxical result does arise if we allow grue/bleen type predicates. But it is possible, within the general framework advocated here, to disallow such predicates. The issue is discussed in *Explaining Science's Success*, pp.70–81.

[10] Paradoxical inductions of this sort are discussed in W. V. Quine and J. S. Ullian *The Web of Belief* (McGraw Hill, 1978).

Premise (1) All fish that we have observed so far have lived before 23rd March, 2017.

All fish live before 23rd March, 2017.

An Eddington inference from Premise (1) might be:

Premise (1) All fish that we have observed so far have lived before 23rd March, 2017.

There exist fish smaller than those we have observed that will live after 23rd March 2017.

But the conclusion of this Eddington inference seems reasonable enough: it is *a priori* unlikely that the time up to which we have observed fish should also happen to coincide with the extermination of all fish. Induction presents us with a puzzle in this case, but Eddington inferences do not.

One familiar puzzle for induction arises from the paradox of the ravens.[11] "All ravens are black" is equivalent to "All non-black things are non-ravens". A positive instance of the latter is a green leaf, but we are very reluctant to say "All ravens are black" receives confirmation from the observation of a green leaf. And so, it is natural to ask: might some counterpart to this paradox arise for Eddington inferences? First let us note that the paradox does not arise for Eddington inferences in any straightforward way. The paradox of the ravens arises only because it is universal generalisations that are (meant to be) confirmed by induction, together with the fact of logic that "All As are Bs" is equivalent to "All non-Bs are non-As". But since Eddington inferences do not confirm universal generalisations, the derivation of the paradoxical result cannot be made.

But still, it might be wondered whether some sort of paradox might be lurking here. The conclusion of an Eddington inference is an existence-statement, for example: There exist fish smaller than four inches. This is logically equivalent to "It is not the case that all fish are four inches or larger". And it might be thought that this is paradoxical. After all, the empirical data on which this conclusion is based is the fact that all observed fish *are* four inches or *longer*. And so, it might appear that an Eddington inference can be represented as:

Inference A:
Premise (i) All *observed* fish are four inches or longer.

It is not the case that *all* fish are four inches or longer.

That is, Eddington inferences lead us to the *negation* of the conclusion to which we would be led by induction. Eddington inferences might, therefore, be thought of as a sort of "counter-induction". And this, surely, is an unacceptable result.

However, it will be argued that if we look at the situation more closely, there is actually nothing implausible or unacceptable going on. First, let us note that as it

[11] See Carl Hempel "Studies in the Logic of Confirmation" *Mind*, **54**, (1945), pp. 1–26 and pp. 97–121.

stands, Inference A is not *quite* an Eddington inference. In order to turn it in to an Eddington inference, it needs to be re-stated as:

Inference A*:
Premise (i*): The blindly chosen size of the holes in our net is four inches, and all the fish in our net are four inches or longer.

It is not the case that all fish are four inches or longer.

Moreover, it seems clear that it would not be a rational use of induction to draw from Premise (i*) the conclusion that all fish are four inches or longer. That this would not be a rational use of induction follows from the justification of induction given in Chap. 2. The justification of induction used there appeals crucially to the nature of the blindly chosen location of our observations. To draw the conclusion that all fish are four inches or longer would be – to return to our example of Chap. 2 – like making an inference from "We blindly chose Geelong as the location from which to observe crows and all the crows there are black" to the conclusion "All crows are black crows located in Geelong". On the view advocated here, if we have blindly chosen some constraint on our observations of Xs – whether that constraint be size, or location, or something else – then induction does not lead us to say all Xs are subject to that constraint; on the contrary, induction permits us to make a generalisation about Xs *independent* of that blindly chosen constraint.

So, in summary, Eddington inferences do not lead us to conclusions incompatible with those to which we are (at least on the view advocated here) led by induction. Provided that the underlying justifications for induction and for Eddington inferences are properly used, the two types of inference lead to mutually consistent conclusions.

Of course, there are some inferences that might be made that *are* clearly irrational. Here is an example:

All fish we have caught in our net have the property of being catchable in our net.

There exist fish too small to be caught in our net that have the property of being catchable in our net.

The inference is clearly absurd, but it is also clear that we are under no pressure to accept it. And the reason for this is because the underlying justification of Eddington inferences given above does not apply in this case. Suppose the conclusion is false; that is, there *do not* exist fish too small to be caught in our net that nonetheless have the property of being catchable in our net. Then there is clearly no sense in which we can say a highly improbable event has occurred. We can be sure the negation of the conclusion is true – that is, that no such fish exist – for the simple reason that the supposition they do exist is logically incoherent. Denying the truth of the conclusion does not lead us to say any highly improbable event has occurred.

In conclusion, the cursory survey carried out in this section would seem to indicate that Eddington inferences would seem to be rather less prone than induction to puzzle and paradox.

5.9 Inference to Molecules

The notion of an Eddington inference has been introduced. Some of the main features of that type of inference have been described, and it has been argued that the notion seems to be able to deal with puzzles and paradoxes rather better than does induction. Let us now consider whether the notion of an Eddington inference might enable us to establish interesting cases of realism. We will consider, more specifically, whether Eddington inferences might be able to establish realism about molecules.

In this section we only consider some general aspects the inference to molecules and other entities postulated to explain observable phenomena. The inferences that were, as a matter of historical fact, used by scientists will be considered in Chap. 7.

Consider the following inference, which we will call Eddington-inference NM:

> All objects capable of being observed by us move around according to Newton's laws of motion. So, by an Eddington inference, there exist objects too small to observe that move around according to Newton's laws of motion.

Note that the above Eddington inference does not merely say "Such objects, if they existed, would obey Newton's laws". Rather, it tells us that such objects (probably) *do exist*. Does this mean that an Eddington-inference thereby establishes that *molecules* exist? This would, I think, be too quick. What a scientific realist wishes to establish, it is surely fair to assume, is that *those entities postulated by a theory of molecules* exist. But how can we be sure the entities to which we are led by the Eddington inference are *the same* entities as those postulated by our theory of molecules? Suppose we postulate molecules (more specifically, tiny material bodies too small to see moving around according to Newton's laws) because they provide us with an explanation of gas laws and Brownian motion. If our concern is with scientific realism, what we want to know is: Are there bodies too small to see that, for example, cause gases to expand when heated? More generally, what we want to know is: Do *certain* unobservable entities, postulated to explain phenomena we have observed, actually exist? More generally do still, do the unobservable entities *postulated by our explanatory theories* exist? The Eddington-inference tells us that *some* entities too small to see, and which move around according to Newton's laws, exist. But are these the very same entities as those that are responsible for phenomena such as pressure increasing with temperature, and the Brownian motion of suspended particles?

5.10 Identifying the Entities to Which We Are Led by Eddington-Inferences with Those Postulated by Explanatory Theories

In this section it will be argued that there is a natural way of establishing that the entities to which we are led by the Eddington-inference NM are the very same as those postulated by at least a simple theory of molecules. First, we need to make the following (oversimplifying) assumption. The assumption is that the entities and properties we can establish by the Eddington-inference NM are sufficient to explain the phenomena we want to explain: the gas laws and Brownian motion. (This assumption *is* certainly at least an oversimplification; this is considered below) But, given this assumption, it seems reasonable to argue in the following way:

> Suppose it were not the case that the entities responsible for the gas laws and Brownian motion were the very same entities as those to which we are led by the Eddington-inference. Then a highly improbable fluke would have occurred: whatever it is that the gas is *actually* made of would have just happened to have the same causal powers as those entities arrived at by Eddington-inference NM, if they were what gases were made of. Since this fluke seems unlikely, we have reason to believe it has not occurred. Therefore, we have reason to believe that the gas *is* made up out of the same entities as those we arrive at by the Eddington-inference NM.

Let us call this last inference the "No coincidental agreement" inference.

The above argument seems, at least *prima facie*, to lead to a form of realism about molecules *without IBE*. There are two steps to the argument:

Step (1): The Eddington inference NM. This gives us probabilistic reason to believe there are unobservable entities that obey Newton's laws.

Step (2): The "no coincidental agreement" inference. This leads us to identify the unobservable entities that obey Newton's laws with the entities responsible for gas laws and Brownian motion.

Both steps are probabilistic inferences: they establish the likelihood of their conclusions, or their greater likelihood than rival hypotheses. Neither, on the face of it, would seem to use IBE. And so, it seems, we have an argument for realism that does not use inference to the best explanation.

In the remainder of this chapter some objections to this suggestion will be considered.

5.11 Objection One: Couldn't IBE Be Recast in Similar Probabilistic Terms?

It might be suggested that we could restate IBE in a way that made it like the inferences defended here. This might be done as follows: Suppose T_1, T_2, ….T_n all explain some observations E. Suppose that T_1 is the best of them. Is T_1 true? Assume it is not, and it is some other theory T_t that is true. If this is so, then – might it not be argued? – a surprising fluke has occurred. The (epistemically) best theory T_1 and the true theory T_t turn out to explain the same observable phenomena E. Since it seems rather unlikely for this to be merely coincidental, we may conclude it is not the case that T_1 and T_t are distinct. That is, we have probabilistic grounds for saying T_1 and T_t are the same, and therefore that T_1 is true.

But, this objection fails. It is, obviously, not improbable or surprising that T_1 and T_t both explain the same phenomena. T_t explains E because it is, by definition, the correct explanation of E. T_1 explains E because it was invented or constructed to do just that. It is the best of the theories we have been able to come up with that explains E. So, clearly, no improbable fluke has occurred.

It is worth briefly considering why the reply just given does not apply against the view defended here. Again, let us assume the true explanation of E is T_t. We find that if gases are composed of the entities to which we are led by the Eddington inference NM they too can account for E. This is a surprising and improbable discovery: an alternative route – an alternative, that is, to looking for an explanation for E – leads us to certain entities, and these also entities turn out to be able to account for E. This is obviously something that is unlikely to be merely due to chance.

5.12 Objection Two: The Argument Given Uses an Unnecessarily Weak IBE-Based Argument for Realism

It has been argued that Eddington inferences can, but IBE cannot, give us probabilistic reason for believing in realism. But, it might be objected, the argument given uses an unnecessarily weak IBE-based argument. More specifically, it might be protested that we do not need to say that the fact that a theory is the best (in the sense of simplest etc.) is, *by itself*, sufficient to justify realism about it. It might (as noted in the previous chapter) be instead suggested that in order for realism with respect to a theory T to be justified, that theory must also have *novel predictive success*. And – it might further be argued – *if* this additional requirement is made, we *do* get a probabilistic argument for realism. This argument is as follows:

> Let T be some theory that makes a novel prediction N, and suppose N is subsequently confirmed. Then (it may be argued) if T were false, it would be highly unlikely that the novel prediction N would be confirmed. But N *has* been confirmed. So, it is probably not the case that T is false. Therefore, T is probably true.________________________(1)

Is this a "purely probabilistic" argument for the probable truth of T? There is a complication. We are supposing that T, in addition to explaining E, successfully predicts – and, we may therefore assume, explains – N. That is, T explains E*, where E* = E&N. But will T be the only theory that explains E*? If it were, then what we have here might perhaps be seen as a case of "inference to the only explanation" rather than "inference to the best explanation". And there is presumably a good deal of plausibility to the idea that if T were the *only* explanation of some phenomena we would have good grounds for supporting it.[12] However, we noted in the previous chapter that if E* is some finite, actually obtained body of data, there is good reason to believe there will be other explanations of E*, even if they are highly complex or *ad hoc*. And we have yet to see what would entitle us to believe T rather than any one of the other theories. The argument given in (1) therefore fails.

5.13 Objection Three: Perhaps the Argument Advocated Here Implicitly Uses IBE

Let us remind ourselves of the argument for molecules sketched here. First, the Eddington-inference NM gives us probabilistic reason for saying that there are entities too small to see moving around according to Newton's laws. It is then observed that if gases were made of these entities, they would give rise to the behaviour described by the gas laws and Brownian motion. This provides us with reason for thinking that the entities to which we are led by the Eddington inference *are* the entities out of which gases are composed. If they were not, then an improbable event would have occurred: whatever it is that gas is actually made of would just happen to have the same causal powers as would those entities arrived at by the Eddington inference NM if they were what gas was made of.

But is this last move in the argument as straightforward as it might seem? It might perhaps be argued that it implicitly uses IBE. More specifically, it might be suggested that we are implicitly assuming that the hypothesis that the two classes of entities – the ones to which we are led by the Eddington inference and the ones postulated by our theory of molecules – are identical provides us with *the best explanation* of the fact that they have the same causal powers. After all, there could be many other explanations of the behaviour of gases, and we have presumably accepted the theory of molecules because we believe it to be the best.

This objection is, however, mistaken. The inference we have here used proceeds as follows: if the entities to which we are led by Eddington-inference NM and the entities responsible for the gas laws were not identical, then an improbable event would have occurred: Whatever it is that gas is actually made of would just happen to have the same causal powers as would those entities arrived at by Eddington inference NM if *they* were what gas was made of. Of course, there may, for all we know, be many types of thing gases are made of, but in the present context that is irrelevant.

One way of bringing out how it is irrelevant is as follows: We do not know what gases are made of – they could be one thing, they could be another. But if gases are not made of the entities to which we are led by the Eddington inference, it seems like a highly improbable fluke that the entities to which are led by the Eddington

[12] Although, as Musgrave has pointed out, this need not be the case if the only explanation we possess is very bad.

inference NM should just happen to have the same causal powers as whatever it is gases are made of. And we have independent evidence – the Eddington inference NM – that entities with those causal powers exist *irrespective* of whatever we may have observed about gases and the movement of, say, pollen seeds suspended in heated oil. This clearly makes it likely that the entities of which gases *are* made are identical with the entities to which we are led by Eddington-inference NM.

We can perhaps put the key point this way. There are, of course, a number of possible ways of explaining gas laws and Brownian motion. But on the view advocated in this paper it is the Eddington-inference, not criteria such as simplicity, which picks out one of the candidate explanations as the preferred, most likely one. And since Eddington-inferences, it has been argued, do make their conclusion more probable, the inference does not appeal to inference the best explanation.

5.14 Objection Four: The View Advocated Here Is at Best Just a Variant on or Special Case of the Argument for Realism from the Concordance of Independent Methods

As we noted in the previous chapter, an influential argument for realism appeals to the concordance of different methods. For example, an influential argument for the existence of molecules appeals to the fact that different methods for determining the value of Avogadro's number yield the same result.[13] But it might perhaps be felt that the proposal advocated here is simply just a variant of this. After all, it might be asserted, the proposal offered here seems to be simply that we ought to say that some sort of theoretical entity exists if we are led to entities of that sort by *both* an explanatory theory and an Eddington inference. In short, the approach advocated here might be described as that of agreement between Eddington inference and explanatory theory.

While there is a sense in which this approach can be seen as a special case of agreement between independent methods, it is important to note it possess a strength not possessed by some other versions of that approach. Let us begin by reminding ourselves of one criticism, made in the previous chapter, of the appeal to agreement of independent methods. If there are several independent methods that tell us the value of Avogadro's number is $N_{A,}$, then we perhaps have good reason to say there are N_A *somethings*, but it is not so clear we have good reason to say that they must be, say, discrete units of matter. There might, for example, only be N_A dispositions to yield a particular experimental result. But it is precisely this difficulty that is overcome by the use of Eddington-inferences. An Eddington inference can give us

[13] Perhaps the first use of the agreement between different methods of determining the value of Avogadro's in defence of scientific realism is Wesley Salmon *Scientific Explanation and the Causal Structure of the World*. (Princeton University Press, 1984).

much more specific information about the type of unobservable entities that exist. If the premise of the Eddington inference is that there exists bits of matter – that is, things possessing the various properties of bits of matter including mass, charge, a location in space and boundaries in space – in a range of sizes down to some small but just perceptible size *d*, then the conclusion of the inference tells us there are probably smaller such entities. That is, the Eddington-inference tells us that entities exist that have at least some of the "nature" generally associated with molecules. The approach advocated here does more than merely tell us there exist N_A somethings: it tells us that those "somethings" are molecules.

In the previous chapter we briefly mentioned the relation the method of concordance had to contrastive confirmation. Concordance, it was argued, would seem to give us good reason to believe a mole contained 6.022×10^{23} molecules *rather than* some other number. But it would not seem to give us reason to believe it contained 6.022×10^{23} molecules, *rather than* entities or causal powers of some other sort. But Eddington inferences do enable us to satisfy this requirement: they do give us reason to believe a mole contains 6.022×10^{23} molecules considered as tiny bits of matter, *rather than*, for example, the same number of dispositions to produce certain observations.

It is also worth noting that the Eddington inference does not merely "make the claim" that there exist tiny, discrete bits of matter. It leads us to this conclusion via a probabilistic inference that seems to be at least as strong as inductive inference. The assertion that there are bits of matter too small to see, when arrived at by an Eddington an inference, would appear to have a degree of evidential support of the same order of strength as that possessed by beliefs such as "The Sun will rise tomorrow". And since that belief would seem to be a paradigm case of a rational belief, so "There are bits of matter too small to see" would also seem to be a clear case of rational belief. Eddington inferences provide us with a "bridgehead" to knowledge about the unobservable realm.

5.15 Objection Five: The Argument Uses an Assumption that Is in Fact False

It is assumed that if gases were composed of the entities to which we are led by Eddington-inference NM, then gases would conform to the gas laws. But, as we have already noted, this assumption is in fact false. We need to attribute additional properties to the entities that make up gases: for example, that collisions between them are elastic. That they are elastic would not seem to be justified by any kind of Eddington-inference from what we have observed.

However, the argument given here still goes through, but in a somewhat weakened form. The assumption that the collisions are elastic is, in the present context, not entirely *ad hoc*. The conjecture being considered is that heat is the random motion of these particles. Collisions involving familiar, observable objects (tennis balls, billiard balls etc.) are not elastic. But (again, oversimplifying somewhat) we can perhaps say in these cases the missing energy is converted in to heat. So, if in the present context we are hypothesising that heat is the motion of the particles making up the gas, and we have (inductive) evidence energy is conserved, then perhaps we do have some kind of reason for believing that collisions between the particles is elastic. (Where else would any energy lost in the collisions of such particles "go"?) If it is allowed that the hypothesis that the collisions are elastic is not *ad hoc*, then our argument would still seem to go through, albeit in a somewhat weakened form. The following claim is surely still unlikely:

> What it is that gas is actually made of just happens to have the same causal powers as would the entities to which we are led by the Eddington-inference NM, together with non-*ad hoc* hypotheses about the nature of such entities, if those entities were what gases were made of.

Note that no assumption is made here that *ad hoc* hypotheses are more likely to be *true*. The role of *ad hoc*-ness is merely to assure unlikelihood of sameness of causal powers.

5.16 Objection Six: The Argument Fails Because a Crucial Inferential Step Is Based on a False Assumption

It might be objected that a crucial step in the argument fails. The first step in the "no coincidental agreement" inference is as follows:

> It is highly unlikely that what gas is actually made of should have the same causal powers as the entities to which we are led by Eddington-inference NM.

That may be true. But what need not be true is the following:

> It is highly unlikely that there should be something, somewhere in the universe, with the same causal powers as the entities to which we are led by the Eddington inference.

In a big universe it presumably *is* likely that there will be something, somewhere, with those powers. So – it may be objected – the argument given here for realism about molecules is not (or might not be) a good one. Suppose we first performed Eddington-inference NM and concluded: "There exists (somewhere) entities with certain causal powers. We then looked around in the universe and finally, after a very long search, found some. Call these entities K. Would we be justified in saying it was highly unlikely entities K had the same causal powers as those to which we were led by the Eddington inference? Possibly not. If we had looked long and hard enough, and it's a big universe, it is perhaps not surprising that we should have found such entities.

However, with respect to the case presently under consideration, this objection clearly "misses the mark". We didn't come across heat and pressure of gases, for example, as a result of some very long "random" search of the things of the universe, and only then note they seemed to have the same properties as the entities to which we were led by the Eddington-inference. In the present cases, there are some general plausibility considerations that "point towards" gases being natural things to be made up out of the entities to which we are led by Eddington-inference NM. The components of liquids and solids seem to be too strongly bound to each other to be moving around like e.g. tiny billiard balls, so if there is anything in our environment made of such things, it seems more likely to be gases. We didn't search around randomly for something that had the right properties, but in this case were led to it by "plausibility considerations".

Of course, this does *not* mean we need to say the "plausibility considerations" increased the likelihood of truth. They simply ensure we weren't looking around for something with the right properties for so long we were bound to find it. Our search was so short our success became surprising.

5.17 A Route to Realism Without IBE

It is suggested that we have a route to realism about some theory T, without IBE, if:

(a) Our explanation T of some phenomena involves postulating entities with properties P.
(b) An Eddington-inference, together with some non-*ad hoc* hypotheses, leads us to assert the existence of some class of entities that have those very same properties P.
(c) The Eddington inference referred to in (b) justifies (according to the position adopted here) belief in the existence of the entities with which it deals.
(d) The "no-coincidental agreement" inference justifies us in asserting the entities to which we are led by the Eddington-inference are identical with the entities postulated by T. Since we are justified in asserting the former exist, we are also justified in asserting the latter exist. So: Realism about T is justified.

5.18 Extending the Scope of Eddington Inferences: Realism about Unobservable Properties

It is natural to object that the scope of the type of realism advocated here is extremely limited. It is, it seems, restricted to entities we may arrive at by some sort of Eddington-inference. And, although these *entities* might be unobservable, their properties, it seems, must be *properties* that are observable.

This topic is discussed in more detail in the final chapter. But here one point will be noted. Let us assume we (somehow) know that observable light is composed of waves, and that red light has the lowest frequency while violet has the highest frequency. We can ask: Are there forms of light with higher or lower frequencies than those we can observe? Suppose there were not. Then, it seems, an improbable event would have occurred: the highest and lowest frequencies that exist would have happened to coincide with what our eyes are capable of perceiving. Since this seems unlikely, we may conclude that higher and lower frequencies probably do exist. And hence that “infra-red” and “ultra-violet” light probably exist.

Here we have, it seems, an Eddington-inference to unobservable *properties*. (e.g., possibly: the property of being “coloured” ultra-violet.) Of course, it may be asked how we know light has a wave character. Perhaps we need IBE to get to know this. But the point of this example is not to show that we can in fact have reason to believe ultra-violet light, for example, exists using *only* Eddington-inferences. It is rather the more modest one of merely arguing that Eddington-inferences may be able to give us reason to believe in unobservable *properties*, as well as entities.

In the final chapter we explore a number of ways in which Eddington inferences can be used to give us reasonable belief about unobservable things other than molecules and atoms. Particular attention is paid to the way they give us reasonable belief about the very large; more specifically, about regions of space and time beyond the observable universe.

Chapter 6
Underdetermination and Theory Preference

Let us begin by reviewing the main results of the previous chapter. It was argued that, provided certain conditions are met, we can construct a purely probabilistic inference to some scientific realist claims. To recap, it was argued that we are justified in adopting realism with respect to some theory T if:

(i) Our explanation T of some phenomena involves postulating non-observable entities with properties P.
(ii) An Eddington-inference, together with some non-*ad hoc* hypotheses, leads us to assert the existence of some class of non-observable entities that have those very same properties P.
(iii) A "no-coincidental agreement" inference justifies us in asserting the non-observable entities to which we are led by the Eddington-inference are the very same as those postulated by our theory T.

However, this account leaves out something that needs to be included in any defence of realism. Consider (i): "Our theory T of some phenomena involves postulating entities with property P." The question naturally arises: Suppose we have, in addition to theory T, some alternative theory T* of the very same phenomena. The thesis of the underdetermination of theory by actual data assures us this may occur. Moreover, suppose that theory T* leads us to assert that there are entities with property Q rather than P. Are we to therefore assert that there exist *both* entities with P and also entities with Q? I take it as uncontroversial this would not be a rationally tenable position. And if we say that we are only to assert the existence of entities with P on the grounds that T is a *better* theory than T*, are we not led back to relying on IBE?

We can illustrate this difficulty with the example used in the previous chapter. There, an Eddington-inference was used which took us from "The observable entities around us obey Newton's laws of motion" to "There are (probably) entities too small for us to see that obey Newton's laws of motion". In the previous chapter the legitimacy of this inferential move was defended, but no argument was presented for the assumption that we do, *in fact*, *have good reason* to believe the observable

J. Wright, *An Epistemic Foundation for Scientific Realism*, Synthese Library 402, https://doi.org/10.1007/978-3-030-02218-1_6

objects around us obey *Newton's laws*. The observable objects around us presumably behave *as if* they obey those laws; but, given underdetermination, their motions could surely be explained by many other possible sets of laws as well. What reason do we have for saying they obey *Newton's laws* rather than some other possible set of laws? If we say that Newton's laws provide the best explanation of the motions and that is why we are justified in saying Newton's laws are true, we are back to relying on IBE.

It is clear that this objection, if sound, would undermine the argument for realism being defended here. As we have defined it, Realism with respect to a theory T is the doctrine that the entities, including the unobservable entities, postulated by T exist and behave more or less as T says they do. Realism about theory T entails that there are entities that (more or less) obey the laws of T. But if we are to be justified in asserting that there are entities that obey the laws of T we need to have good reason to assert they obey the laws *of T* rather than some other theory that can explain the same data. In the absence of such a justification, realism *with respect to T* would not seem to be justified.

It might perhaps be feared that the obstacle with which are now confronted is insuperable. In order to justify realism with respect theory T, we need to have good reason to prefer T to other theories that explain the same phenomena. The most natural way of doing this – perhaps the only way – is to say T is preferable because it is the best explanation. But Chap. 5 was devoted to arguing that we have yet to see how an explanation's being the best constitutes good reason for saying it is true or probably true. We seem, therefore, to be without any way of justifying realism with respect to any theory T.

However, the situation is perhaps not as hopeless as it might at first seem. First, let us remind ourselves of exactly why IBE seems to fall down as a way of establishing realism. The fact that an explanation is the best does not (it has been argued) give us good reason for saying it is true *at the theoretical* level. In Chap. 3 we noted that the history of science seems to show that the inference from "T is our best explanation" to "The *unobservable entities* postulated by T" seems to be especially risky. However, we might yet have, or be able to construct, good reason to believe at least what a theory says at the *observational* level. In fact, in Chap. 5, it was noted that perhaps we do have this in a limited class of cases. For example, we noted that there did seem to be (simplicity-related) criteria that gave us probabilistic reasons to prefer certain "smooth" curves drawn through data points to other possible curves. If the data points are only about *observable* properties, then these (simplicity-related) criteria may, it seems, give us good reason to believe that some theories about observables were more likely to be true than other rival theories.

This suggests a possible strategy for developing our argument for scientific realism: First, we develop criteria for showing one theory to be more likely than another *as an account of the behaviour of observables.* Armed with such as an account, we might get good grounds for saying observable entities conform to certain regularities. Call these regularities R. We then use an Eddington inference to derive the conclusion that there exist *non-observational* entities that conform to R. Finally, the No-coincidental Agreement Inference then permits us to conclude that the entities to which are led by the Eddington inference are the very same as those postulated by our explanatory theory T.

It is useful to illustrate this with an example. We see objects in the world around us in motion: balls rolling down slopes, pendulums, pucks sliding across ice and so on. To what laws or regularities are these objects conforming? Finding out to what laws or regularities *these* objects are conforming need not involve postulating any unobservable entities. It *may* involve nothing more than something like finding the smoothest curves that can account for their motions, and extrapolating or generalising from them. And, *perhaps*, there are purely probabilistic modes of inference that can do that. Now, let us suppose we can do this and succeed in giving a probabilistic justification for the claim that these objects are conforming to Newton's laws of motion. So, we will have succeeded in showing that at least to within the limits of observational accuracy *objects large enough to be observed* (probably) conform to Newton's laws. Then, an Eddington inference carries us to the conclusion that there (probably) exist *unobservable* entities that also conform to Newton's laws. We can also show that if such objects existed they would, at the macroscopically observable level, behave in the same way as heated gases. But we now note that according to the kinetic theory of heat there are unobservable objects, obeying Newton's laws, that are responsible for the observable phenomena associated with heat. The No-Coincidental Agreement inference entitles us to draw the conclusion that the entities postulated by the kinetic theory are the same as those to which we are led by an Eddington inference, and hence that the former do in fact probably exist.

Plainly, the above is based on the assumption that we can have good, probabilistic reasons for saying that the objects that we observe – that is, objects such as rolling balls, sliding pucks etc. – conform to Newton's laws. More generally, it is based on the assumption that we can have probabilistic reason for saying observed entities conform to one set of laws or regularities than another. But, if theory is underdetermined by actual data, it is not evident that we can have good reason for this.

The aim of the present chapter is to defend the thesis that, actually, we can have such good reason.

6.1 Illustration: A (Very) Brief Sketch of the History Astronomy

The ideas outlined in the remainder of this chapter are largely a summary of earlier work.[1] Some key ideas are not as precisely defined as they might be, and numerous natural objections are not considered. These things have, however, been done elsewhere.[2]

[1] See Wright *Science and the Theory of Rationality* (Avebury, 1991) and *Explaining Science's Success*: *Understanding How Scientific Knowledge Works* (Acumen, 2013).

[2] See Wright (1991, 2013).

We are presently concerned with the question: What makes one way of discerning a pattern or order on some phenomena better than another? A good place to start is by comparing some very broad brush sketches of episodes from the history of thought about the solar system. It must be emphasised that the following discussion is *not* in any sense intended to be history of science. *It is, rather, an exercise in taking some familiar ideas from the history of astronomy to illustrate a conceptual point.*

What form might an account of our observations of the night sky take? One possible "account" might just be a *record* of the positions of all the heavenly bodies from night to night, and from time to time on a single night. Such a record might include entries such as: "At 9.10 pm, such and such a star was at such and such a position in the sky". Similar entries would be given for all stars at all times on all evenings. Such a list would, of course, be immensely long. It would, by itself, not furnish us with any *predictions* concerning where any heavenly body would be. And neither would it furnish us with us with any kind of *explanation* of what was going on. It would just be a record of the data. But still, we will say it would qualify as "an account", of sorts, of astronomical phenomena. We will refer to it here as "A1".

Another account of the solar system might be based on the work of Aristotle, Ptolemy and Apollonius of Perga. On such a model, the Earth is a sphere, lying at the centre of the universe. Around the Earth is a series of concentric spheres, made of a transparent substance, such as crystal. The number of these concentric spheres might vary. In some models actually advanced, the spheres were all rotating around the Earth at different rates. In one early version of the model the heavenly bodies were embedded in one or another of these spheres, and so traced a circular path around the Earth.

Of course, models of this sort still need to be able to account for the retrograde motion of the planets. And to account for this the notion of "epicycles" was introduced. Any given planet was thought to also be moving in another, smaller, circle around a point embedded in a sphere itself moving around the Earth. On this view, any given planet was moving in two circles. First, it was moving in a large circle around the Earth. But second, it was also moving in a smaller circle around a point itself moving in the larger circle. The smaller circle was the "epicycle".

The idea of epicycles provided a possible explanation for retrograde motion. Consider what would be observed from Earth as a planet moved both around the Earth in a larger circle, but also moved in a smaller epicycle. When the motion of the planet around the Earth in the larger circle, and in the epicycle, were in the same direction, the planet would seem to be moving in that direction rather quickly – at least relative to its velocity at other times. But as the planet moved around the epicycle, and the direction of its motion in the epicycle became perpendicular to the direction of its motion in the larger circle, the planet, as viewed from Earth, would appear to have slowed down. As the planet continued around the epicycle, the direction of its motion around the epicycle would become the opposite of its direction of motion in the larger cycle. If it was travelling around the epicycle faster than the larger circle, there would come a point at which, viewed from here on Earth, the planet would seem to be moving backwards. Then, as it continued, its motion in the epicycle would again become perpendicular to its motion in the larger circle: it would seem to have returned to its usual direction of motion. And so: the hypothesis of epicycles would seem to be able to account for at least the general features of retrograde motion.

What was impressive about the hypothesis of epicycles was not merely that it was able to account for the *general features* of retrograde motion, that is: motion in one direction, slowing, a brief period of reversed motion, and then return to the usual direction – it was also able to account, with considerable precision, for the rate at which planets slowed, reversed and then resumed their path.[3] On the face of it, we are perhaps inclined to say that it is *unlikely to be merely due to chance* that the rates at which planets slowed, reversed and then resumed their path should correspond so closely to that which would be observed if they were following epicycles. This would seem to lend at least a measure of plausibility to the hypothesis that *they are*, in fact, moving in epicycles.

Unfortunately, however, closer observation revealed that this picture was not quite right. It was originally thought that the Earth lay at the centre of the larger circles around which the centres of the epicycles moved. However, to accommodate actually observed data, it was found necessary to have the centres of the epicycles moving around a point shifted some distance from the position of the Earth. Even then, the observed data did not quite fit the predictions of the model. It was found closer fit could be obtained by adopting a still more complex system. According to this more complex system, some planets were said to be moving in an epicycle (or, rather, an "epi-epicycle") the centre of which was a point itself moving in an epicycle. This was a feature of the Ptolemaic model of the universe. Although highly complex, the Ptolemaic model was able, with a high degree of accuracy, to fit the observed data. We will refer to this model as "A2".

The third model of the solar system we will consider is that devised by Kepler. Following Copernicus, Kepler placed the Sun rather than the Earth at, more or less, the centre of the solar system. All the planets, including the Earth, rotated around the Sun in an ellipse, with the Sun at one focus. In addition, the Earth rotated once on its axis every 24 hours, giving the appearance of the heavenly bodies rotating around the Earth in circles. We will refer to this model as "A3".

We will not here be concerned to evaluate all the arguments that were raised for or against A2 or A3. There are, in fact, some respects in which the model of crystalline spheres would appear to have an advantage over the Keplerean model. The idea that all the heavenly bodies were embedded in crystalline spheres provides a natural mechanical explanation of why everything was rotating in circles and epicycles. But the possibility of this mechanical explanation is lost if it is asserted the planets are moving in ellipses rather than circles.[4] We will only be concerned with comparing

[3] For a discussion of the impressive degree of accuracy of Ptolemy's system, see "Contra-Copernicus: A Critical Re-estimation of the Mathematical Planetary Theory of Ptolemy, Copernicus and Kepler" by Derek J. de S. Price in *Critical Problems in the History of Science* edited by Marshall Claggett, (University of Wisconsin Press, 1959), pp. 197–218.

[4] It is again perhaps worth stressing that this account is not offered as serious history of science. But its aim is not to provide an accurate account of an episode from the history of

A2 and A3 as competing ways of *discerning patterns in the data*, and not with a mechanical explanation of the patterns discerned. A1 provides us with the data, A2 and A3 provide different ways of describing the patterns in the data.

First let us note that there are a number of respects in which there seems to be something impressive about A2. A1contains a vast mass of data. We are perhaps inclined to overlook just how *a priori* unlikely it should be that such a vast mass of data should come even close to being explainable by such a simple idea as "Everything is moving in circles around the Earth." Of course, this very simple idea was found to not be satisfactory, but the fact that such a large quantity of data comes so close to conforming to such a simple pattern might perhaps persuade us there is something right about it.

A2 can be seen as a development of the idea that everything is moving in circles, but not necessarily around the Earth. We have already noted that A2 makes use of "epicycles" to explain retrograde motion. Also, epicycles do not merely explain the *broad* features of retrograde motion: motion in one direction, slowing, reversal, and then resumption of motion in the original direction. The hypothesis was also able to explain the observed motion of planets exhibiting retrograde motion to an impressive degree of accuracy. And so, the bold hypothesis that everything was moving in a circular path was able, with a surprising degree of accuracy, to explain the motion of bodies which, at least on the face of it, did not appear to be moving in circular paths at all. Confronted with this surprising and *a priori* unlikely match between such a bold and simple idea and a large mass of observed data, we are perhaps inclined to say: "This couldn't be merely due to chance."

But, of course, A2 also has features that make it less convincing. Notoriously, the hypothesis of epicycles required more and more *ad hoc* modification to get it to fit

---

science. The aim, rather, is to use some familiar concepts from the history of science to illustrate a *conceptual* point.

It would, of course, have been desirable if there had been some episode from the history of science that illustrated the conceptual points in a way that could be stated very briefly. Unfortunately, however, I have not been able to find any such episodes. The example of the Ptolemy versus Kepler is the nearest I have been able to come to something that illustrates the conceptual points reasonably simply.

I have elsewhere argued that a number of episodes from the history of science do exemplify the concepts illustrated here. These episodes are: Newton's argument for universal gravitation, the transition from the phlogiston theory of combustion to the oxygen theory of combustion, Einstein's development of the Special Theory of Relativity and Mendel's development of genetics.

It might perhaps be suggested that the absence of any examples in the history science that simply and briefly exemplify the concepts is a sign that these concepts are not really those used by scientists in evaluating theories. However, it seems to me that this conclusion does not necessarily follow. The real world is often extremely complex. A leaf being blown the street might not obviously be conforming to Newton's laws of motion at all, but to a very high degree of approximation, it is. It has been argued that in the examples cited above, scientists really do prefer those theories that maximize the independence of theory from data, but showing *how* they do often involves a lot of detail. One reason for this is because the chains of reasoning used by scientists are often very complex. I have argued that each individual (ampliative, non-deductive) step in their reasoning involves choosing the most independent theory, but since there are many such steps in their reasoning, the overall picture can be quite complex.

the data. As we have noted, the centre of the larger circle in which some planets were thought to be moving had, in an *ad hoc* manner, to be moved from the location of the Earth. And, in some cases, “epi-epicycles” had to be added to epicycles to get the model to fit the observed data.

Let us now consider A3. The most obvious contrast between A2 and A3 is the much greater simplicity of the latter. Instead of the large number of epicycles required by A2, A3 simply says that each planet is moving in an ellipse, with the Sun at one focus, and the apparent circular motion of the stars is due to the Earth rotating on its axis. Intuitively, it seems enormously *a priori* unlikely that the huge mass of information in A1 should be capable of being explained by such a simple set of hypotheses. We are perhaps inclined to reason: it is highly unlikely that it should merely be due to chance that the data in A1 exemplifies such a simple pattern, therefore, it seems likely that it is *not* merely due to chance, and hence – it seems likely – A3 has got on to something at least close to the laws that actually govern the motions of the heavenly bodies.

In summary, the inference leading us to the preferability of A3 is something like the following: it is *a priori* highly unlikely that such a vast mass of data should, by chance, be explainable by such a bold and simple hypothesis, so it is likely it is not merely due to chance; Therefore: it is likely A3 has got on to something like the actual laws governing the data. This inference is rather more compelling in the case of A3 than it is with A2. A2, we recall, was a complex structure of epicycles upon epicycles. On the face of it, it perhaps seems rather likely that we could get an object to follow, with a fair degree of closeness, *any* curve at all if we were to make the system of epicycles upon epicycles complex enough. We can always get a system of hypotheses to fit the data with enough *ad hoc* tinkering. But still, there is something impressive about A2. It was, for example, able to explain with an impressive degree of accuracy some cases of retrograde motion by postulating *just one* epicycle. And *a priori* it seems unlikely that the puzzling motion could have been explained by such a simple hypothesis.

Let us now describe in more general terms why we feel than A3 is better than A2. First we will begin by noting an obvious point: it is always possible to get “an account”, of sorts, of any body of data by simply giving the data themselves. This is what A1 does. There is, presumably, no *a priori* reason why any arbitrary body of data *must* be capable of being explained by a theory simpler than the body of data itself. But it also seems likely that, for any large body of data, there will be some sort of simple pattern exhibited somewhere in it “by chance”. For example, perhaps the numbers generated by some random-number device might produce “by chance” the sequence 0, 1, 2,….,9. But sometimes we might find in a body of data some strikingly simple, or surprising, pattern of which we are inclined to say “This couldn’t be here simply due to chance!”. And A2, in some places, identifies such a pattern: we are perhaps inclined to say it could not merely be due to chance that the retrograde motion of the planets could be explained with such a degree of accuracy by the hypothesis of epicycles. More generally: the larger the body of data, or the more strikingly bold or simple the hypothesis capable of explaining it, the more we are inclined to say that it could not be merely due to chance that the data is explainable by this hypothesis. Also: the less likely it is that it is merely due to chance that the data is explainable by the hypothesis, the more likely it is that the hypothesis has got on to some genuine tendency or propensity that exists in nature for the data to conform to that hypothesis.

We need a term for the type of theory of which we are inclined to say “It can’t be just due to chance that the data conforms to this theory.” The term to be used here is “the *independence* of theory from data”. If a theory exhibits a great deal of *ad hoc dependence* on the data then it is hardly surprising that it should fit the data. But if it does not exhibit *ad hoc* dependence – if it is highly *independent* – then it is surprising that it should be able to fit the data, and so we are inclined to say “This couldn’t be due to chance.”

In our discussion of the history of astronomy, it is clear A1 exhibits the greatest degree of *ad hoc* dependence on the data: it is simply a record of the data, and each and every aspect of it depends upon the data obtained. A1, we can say, achieves no independence from the data at all. It has the lowest possible degree of independence. A2 does achieve a degree of independence. We have already referred to its treatment of retrograde motion. But A2 also required an *ad hoc* proliferation of epicycles to achieve consistency with the observed data. A3, by contrast, has less *ad hoc* dependence on the data than either A2 or A1. Of our three examples, A3 has the greatest degree of *independence*.

We are now in a position to give a (*very* rough) definition of the notion of the independence of theory from data. Independence is a property of theories that explain some body of data. The greater the number of “aspects” or “components” the theory has, the greater its degree of dependence on the data. Conversely, the smaller the number of such components, the greater its independence. So, if two theories explain *the same* body of data, the one with the smaller number of explanatory components will be the more independent. However, if two theories with the same number of dependent explanatory components explain different bodies of data, the theory that explains the smaller number of components of data has the higher degree of dependence. The higher the ratio of dependent explanatory components of theory, the higher the degree of dependence of a theory. And since independence is the inverse of dependence, we may define the independence of a theory as the ratio of the number of components of data it explains to the number of dependent explanatory components of the theory.

$$\text{Independence of T} = \frac{\text{Number of components of data}}{\text{Number of dependent explanatory components of theory}}$$

This definition leaves it unexplained what is to count as a “single component of data” or “dependent explanatory component of theory”. Some remarks on this matter are given later in this chapter, but the author has gone in to the topic in greater detail elsewhere.[5]

[5] See the author’s Science and the Theory of Rationality (Avebury, 1991) and *Explaining Science’s Success: Understanding How Scientific Knowledge Works* (Routledge, 2013).

We are now able to give a rough sketch of why independence should be a desirable property of theories. Pretty clearly, greater the amount of *ad hoc* tinkering to which we are prepared to subject our theory, the less surprising it should be that the resultant theory can accommodate the data. On the other hand, the more independent a theory is from the data – roughly: the bolder or simpler or smaller the number of dependent components the theory has – the more surprising or *a priori* unlikely it is that the data should be able to be explained by such a theory. We may therefore say: the more independent from the data is some theory that can account for the data, the less likely it is that it should merely be due to chance that the data should conform to a theory with that degree of independence. Hence: the more likely it is that it is *not* merely due to chance that the data should be conforming to a theory with that degree of independence. Consequently, the more likely it is that there is a tendency or propensity for the data to conform to that theory. And: if it is likely that there is a tendency or propensity for the data to conform to the theory, we may reasonably expect future data to conform to the theory. So: theories with a high degree of independence are more likely to lead to successful prediction.[6]

6.2 Conformity by Data to a Theory “by Chance”

In the previous section we made reference to a body of data conforming “by chance” to a theory. It is worth getting a little clearer on this notion.

Let us suppose that scientists are observing an object O. The object O has a property P, and the magnitude of P changes with time. Scientists observe the values of P at different times and record their observations. Their observations, we will say, take the form: “At time t_1 the value of P was n”, “At time t_2 the value of P was n*” and so on. The class of all possible such observations of the values of P for O will be very large, and may be infinite. Let us call the class of All such Possible Observations APO.[7] Scientists, of course, will only be able to obtain a fraction – perhaps a very small fraction – of the observations in APO. Of all the possible observations in APO, which ones will be the ones that scientists actually obtain? This, will, of course, be determined by a variety of contingent factors. It will, for example, be determined by the date on which they happen to commence making observations. It will be determined by the date at which they cease making observations, and by how

[6] This is covered in more detail in Wright (1991, 2013).

[7] We will not here attempt to explicate the notion of “all possible observations” precisely, but a rough way of doing it is as follows. What we are after, in our example, is the set of all possible true values of the observable property P at different times. Let us represent any such a statement as a triple <P, t, n>. This tells us that at t property P has value n. Now, let W be the set of possible states of the world in which the laws of nature are the same as they are in the actual world, but in which the “initial conditions” may vary. We will say that a <P, t, n> is a possible observation if and only if there exists a world in W in which <P, t, n> is true. And the set of all possible observations is just the set of all such true-in-some-member-of-W possible observations.

frequently they happen to make their observations. Their observations, we will say, will have a *location* in APO.

The aim of the scientists, we may assume, is to discern some pattern in the data. But: in advance of making observations, they do not know what the pattern is, or even if there is a pattern there at all. Also: as far as they know, some parts of APO might exhibit one pattern while other parts exhibit another. Or, some parts might exhibit a pattern while others exhibit none at all. But, in advance of making their observations, scientists do not know which parts may, or may not exhibit a pattern or patterns. In this state of ignorance, scientists nonetheless make their observations of some part(s) of APO. Their observations, we will say, have a *blindly chosen location* within APO.

Now, let us suppose, scientists make only three readings. On the first reading, the value of P is n, on the second it is n + 1, and on the third it is n + 2. The value of P has exhibited a simple pattern: it has increased by one each time. It is natural to wonder whether we have here got on to a genuine rule or regularity in the behaviour of the value of P, or whether its observed conformity to this rule is just "a fluke". And a natural way of clarifying this question is: Would the values of P also conform to this rule if we were to obtain readings from other locations in APO? If the data would continue to conform to this pattern, then we can say that the fact that the data exemplifies this pattern is not merely due to the blindly chosen location of the data.

It has elsewhere been argued that the greater the degree of independence a theory has, the less likely it is that it is merely due to the blindly chosen location of the data that the data should conform to a theory of this degree of independence.[8] Therefore, the higher the degree of independence of a theory, the more likely it is that our data would conform to the theory if the data had been obtained from some other (blindly chosen) location. Hence, the more independent a theory is from the data, the more likely it is that future observations will continue to conform to the theory. The more independent a theory is, the more likely it is that it will continue to be empirically satisfactory.

6.3 Replies to Criticisms

Jarrett Leplin has objected to this argument for the preferability of highly independent theories on the grounds that it is unnecessary.[9] He asks what point there is developing such a *justification or argument* for independence rather than merely accepting it as an *a priori* plausible but defeasible claim that nature is simple.

One very general reply is that a justification of a preference for highly independent theories is required for the same reason we generally require arguments in philosophy or proofs in mathematics. A claim P is advanced that may be plausible

---

[8] *Explaining Science's Success*, Chap. 4, pp. 57–94.

[9] See Jarrett Leplin "Book Review: Explaining Science's Success" in *Analysis*, v 74, (2014), pp. 184–185.

to some, but more questionable to others. Or, it might have some degree of *a priori* plausibility but that degree of plausibility might not be very high. The aim of an *argument* or *justification* is to show how P can be derived from other claims that have a higher degree of *a priori* plausibility than P itself, or are less disputed than P.

There are, however, some more specific replies.

First, note that the defence of independence, like the defence of induction, relies upon or "stands upon" the argument, given earlier, for the thesis that we can have good *a priori* reason for some synthetic propositions. This is certainly non-redundant: unless such a thesis is accepted the defence of independence may be rejected "on principle".

Second, the conclusion that independent theories have an increased chance of empirical success *is derived from* premises or assumptions that are less tendentious than the baldly made claim "independent theories have an increased chance of success". It is derived from the claim that bodies of data that exemplify patterns with a high degree of independence are rarer than those that do not. And this claim was itself argued for by appeal to some basic features of probability.[10] The defence of independence has a feature we want from any argument that purports to be rationally persuasive: it shows how the more tendentious or dubious can be derived from that which is less tendentious or dubious.

Third, it was also argued that the approach can naturally deal with cases that are generally troublesome for defences of simplicity, such as "grue" and the fact that simplicity can be relative to mode of representation.[11]

Matthias Egg criticises the notion of a propensity used in the argument.[12] Briefly, Egg claims that in order for us to be justified in saying there is a propensity for As to be Bs, we must first know something about "the causal structure of the world". Suppose we have observed many crows and they have all been black. Then, we may with considerable plausibility assert that crows have a propensity to be black. But we would not be justified in saying black things have a propensity to be crows. Moreover, we would not be justified in saying this even if (improbably) crows were the *only* things we had observed that were black. And the reason for this is because we know that blackness is *not the type of property* that could cause a thing to be a crow. We are justified in saying that there is a propensity for crows to be black since we know something's being a crow *could* be causally responsible for it being black, and we are not justified in saying there is a propensity for black things to be crows because we know, or believe, blackness *could not* cause a thing to be a crow. More generally, Egg argues, in order for us to be justified in saying there is a propensity for As to be Bs, we need to already be in possession of knowledge of "how the world works" or, as Egg puts it, the "causal structure of the world".

It seems to me that Egg somewhat overstates the extent of the difficulty presented by this type of consideration. There is of course some truth to the point he

[10] For more detail, *Explaining Science's Success*, p. 68.

[11] See *Explaining Science's Success*, pp. 70–81.

[12] See Matthias Egg "Book Review: Explaining Science's Success" in *Dialectica*, v. 67 (2013), pp. 367–372.

makes. For example, given that we know that crowness is the type of property that could cause blackness, the observation of just a few black crows makes us pretty confident that crows have a propensity to be black. Also: given that we know, or believe, that blackness is not the type of property that could cause crowness, the observation of even a very large number of black things that are crows would not incline us to believe that black things have a propensity to be crows. But: it need not follow that if we *do not know* whether blackness could cause crowness, or vice-versa, that we are thereby prevented from having good reason to make any claim about propensities.

It can be useful, I think, to approach this matter with the following hypothetical case, loosely based on the type of star known as a "Cepheid variable". Suppose we discovered that the brightness of certain stars increased, decreased and then increased again, regularly. For some of these stars the time between these pulses might be a few days, for others it might be a number of months. Now, suppose that initially we do not know what sort of causal relationship there might be between the periods of brightness and the periods of dimness. We do not know, for example, whether a period of brightness caused a period of dimness, which in turn caused another period of brightness, or whether the periods of brightness and dimness were themselves both epiphenomena of processes occurring within the star. We do not know about the "causal structure of the world" with respect to stars of this type. But still, we would only need to observe a fairly small number of periods of dimness and brightness to become persuaded that stars of this sort had a *propensity* to exhibit alternating periods of brightness and dimness. Knowledge of the causal structure of the world is not *necessary* for it to be rational to believe in propensities.

There is a related matter that concerns Egg. He suggests it is somehow inappropriate to speak of *data* having a disposition to conform to certain patterns. And I can certainly agree it would be very strange to say that earlier data had a causal power to produce later data. But: to say that a propensity *exists* is not necessarily to *locate the causal mechanism* responsible for that propensity. To stick with our example, to say that *there exists* a propensity for certain stars to have alternating periods of brightness and dimness is to not to say it is the brightness itself that has the causal power to produce the dimness. Likewise, to say there exists a propensity for data to conform to a particular pattern need not be to assert that it is the earlier stages of the data that are causally responsible for the later stages conforming to the pattern.

Mario Alai has argued that the view advocated here requires a realist ontology.[13] He says that if there exists a propensity for some data to conform to some pattern, then there must exist some mechanism causally responsible for the conformity to the pattern. Thus, the view advocated here requires us to say such mechanisms exist. And this, if I understand Alai correctly, is tantamount to realism.

Of course, there are many different doctrines, of varying degrees of strength, that can reasonably be called forms of realism. But to merely say that there exists *some* mechanism that is responsible for our data conforming to theory T is surely very

[13] See Mario Alai "Book Review: Explaining Science's Success" in *Metascience*, v. 23 (2014), pp. 125–130.

different from saying that the unobservable entities postulated by T exist and behave more or less as T says they do. For example, to say that there is a disposition for the data to conform to our theory of electrons is very different from saying that electrons exist and behave (more or less) as our theory says they do. The mere assertion that there is a disposition for our data to behave in this way is logically weaker, in the sense of asserting less, than realism with respect to our theory of electrons. Of course, in one sense of the word "realist", saying there exists a propensity for our data to conform to this pattern, or that there exists *some* mechanism responsible for this conformity, is "realist". It is to say that there exists *something* more than the data themselves. But it is surely logically weaker than *realism about electrons*, for example. So, while reference to "propensities" is in some sense "realist", it clearly logically weaker than *scientific* realism.

If I understand Alai's objection correctly, it is that in asserting that there exists a propensity for the data to conform to a particular pattern, or that there exists some mechanism responsible for this conformity, I am availing myself of a more generous ontology than I am entitled to. *If* this is Alai's point, then it seems to me to be based on a possible misconstrual of my aims. My aim, in the work Alai criticises, was to explain the ability of scientists to hit upon theories that subsequently enjoy novel predictive success. I argued that to explain this by saying that the theories that scientists use to make these novel predictions are *true* only presents us with another problem: how have scientists managed to hit upon theories that are *true*? This problem would seem to be, if anything, more difficult than the original problem of explaining how they hit upon theories that went on to have novel predictive success. With respect to the problem of explaining how scientists have hit upon theories with novel success, this proposed solution would seem to take us backwards. My aim, in the work Alai criticises, was to offer an explanation of how scientists hit upon a successful theory without attributing to them knowledge of the truth of the theory. But, of course, none of this means that realism is not *true*. Neither does it mean we cannot have good reasons for realism. Asserting that the move replaces an already difficult problem with an even more difficult one does not commit us to a non-realist ontology.

It might be helpful to return to our earlier example, loosely based Cepheid variables. Suppose a scientist has observed such a variable star. We may assume that the scientist has observed that the period of the variations in luminosity is one week. We will the suppose that the brightness is recorded by a pen making a line on a moving graph paper. When the brightness is high, the line moves up, when it is low, the line moves down. We can regard this line as "the data". The scientist has observed the brightness going up and down each week for a large number of weeks. On the view advocated here, it is reasonable to suppose there exists a propensity for the star to conform to this pattern. It is also, I think, harmless enough to say there is a propensity for *the data* (in this case, the line) to conform to the pattern, provided there is no implication that the mechanism causally responsible for conformity to the pattern is located (entirely) within the data itself.

Now, as noted, it is one thing to say there exists (presumably within the star itself) *some* mechanism, the nature of which we may be ignorant, that is responsible for this propensity, but quite another thing to say that we, or anyone, knows *precisely what that mechanism is*. And, of course, there are many problems concerning how we might know *precisely* what that mechanism is: those problems are the concern of the present book. Moreover, to say that scientists successfully made novel predictions that were subsequently confirmed *because* they derived the predictions from the true account of that mechanism seems to present us with a phenomenon even more puzzling than that which it is meant to explain. However, the logically and epistemically much weaker assertion that there exists *some* mechanism (unknown to us) which is responsible for the propensity is confronted neither with that problem, nor any of the other epistemic problems associated with specific realist theses. As far as I can see, there is nothing in the view advocated here, or in the work Alai criticises, that precludes me from saying that *some* such mechanism exists.

6.4 Realism and the Notion of Independence

In the previous section we noted that a high degree of independence increases the chances of subsequent empirical success. We should observe, however, that there is no reason to suppose that theories with a higher degree of independence are more likely to be *true*, in the sense of "true" with which the scientific realist is concerned. Let us assume that we can, roughly speaking, divide a theory up in to its purely observational component, and its components that have something to say about non-observational, theoretical entities. We say a theory is true *tout court* if all of its theoretical *and* observational components are true. Nothing in the above argument gives us any reason to assert that *both* the observational and theoretical components of a highly independent theory are true: it provides us with no reason for saying such a theory is true *tout court.*[14]

[14] These remarks presuppose that *there is* a distinction between the "observational level" and the "theoretical level". Such a distinction has of course proved difficult to precisely define. However, it is skepticism about scientific realism that is surely based on the assumption that some such distinction exists, even if it is "fuzzy". A skeptic about scientific realism will presumably advance theses such as: "Some claims are clearly *less* observational than others, as claims become *less* observational their epistemic status becomes *more* dubious, and many of the statements of contemporary science are so far removed from the observational level, and have such a low epistemic status, it is not rational to believe them." It is the last of these claims, it will here be assumed, that scientific realists are out to refute.

6.5 The Independence of Theory from Data and Popperean Boldness

There are important relations between the notion of independence of the data and the Popperean notion of "boldness".[15] A "bold" hypothesis will, generally, in some sense be "based on" empirical data but it will go beyond that data: it "sticks its neck out" and makes a risky claim about how things are beyond the data; and the riskier the claim, the bolder it is. In this rough, intuitive sense, the bolder a hypothesis, the more independent of the data it is. For Popper, the bolder or more falsifiable a hypothesis, the less *a priori* probable it is.[16] And on the view advocated here, the more independent a theory is from the data, the less *a priori* likely it is that, through chance, the data should exemplify the theory. But there is also an important respect in which the view offered here differs from that of Popper. On this view, highly independent theories are less likely, by chance, to be exemplified by data; but, *if* the data *does* exemplify a highly independent theory, it is more likely that this is *not* due to chance and hence that there is a propensity for the data to exemplify the theory. And if that is so, it is more likely that the data will continue to exemplify the theory. Highly independent theories, although *less* likely to be exemplified by the data, if they are so exemplified, are *more* likely to enjoy subsequent empirical success.

6.6 Summary of the Argument for the Preferability of Highly Independent Theories

It is useful to have before us a summary of the argument for the thesis that the independence of a theory increases its chances of empirical success.

(1) The more independent a theory is from the data, the less likely it is that it should be merely due to the blindly chosen location of the data that the data exemplify a theory of that degree of independence.
(2) The more independent a theory is from the data, the more likely it is that it is not merely due to the blindly chosen location of the data that the data conform to the theory.
(3) The more independent a theory is from the data, the more likely it is that there is a propensity for the data to conform to the theory.
(4) The more independent a theory is from the data, the more likely it is the data would conform to that theory if the data were to be obtained from other (including future) locations.

[15] Popper uses the term "boldness" to describe what he regards as the central desirable quality of scientific theories in *The Logic of Scientific Discovery* (Routledge, London, 2002), p. 280. The similarity between Popper's view and the notion of the independence of theory from data is noted in Brad K. Wray's review of Wright (2014) *Australasian Journal of Philosophy*.

[16] See for example Popper, Karl *Conjectures and Refutations* (Routledge, London, 1963), p. 315.

The author has defended this argument elsewhere.[17] For our present purposes, the thing to note about the argument is its *probabilistic* character. If sound, the argument

[17] See Wright, J *Science and the Theory of Rationality* (Avebury, 1991) and *Explaining Science's Success: Understanding How Scientific Knowledge Works* (Acumen, 2013), esp. pp. 66.

It is worth looking a little more closely at just what it is the argument aims to establish. The conclusion of the argument, as presented here is: "(4): The more independent a theory is from the data, the more likely it is the data would conform to that theory if the data were to be obtained from other (including future) locations." It is claimed that the argument that has been given for (4) makes (4) *epistemically probable*.

Objections have been raised against the adequacy of saying that (4) is (merely) epistemically probable. (See, for example, K. Brad Wray's review in Australasian Journal of Philosophy 91 (4), pp. 833–834. (2013). In *Explaining Science's Success* I offered the notion of the independence of theory from data as a way of explaining certain types of novel predictive success in science. Moreover, I acknowledge that, at least on the face of it, appealing to purely epistemic probabilities to explain certain events, such as empirical successes, might seem to be problematic. But here the episemic probability of (4) is not used to explain any subsequent events. It is rather used to show what makes a belief rationally acceptable, or rationally preferable to others.

However, is it really the case that it is not legitimate to use (4) to explain the success of science? It will be argued that actually it is perfectly legitimate. This will become clear, it will be argued, if we attend to the distinction between an explanans *having* an epistemic probability, and an explanans being the assertion of an epistemic probability.

Let us begin with a simple example. Suppose a dice is tossed a large number of times, and it is observed to come up "six" about half the time. Then, we will surely be inclined to say that the dice is probably unfair, or weighted. Assume that, on the basis of our observations, we conclude:

The propensity for a six to come up is about one half.________________(PS)

Although PS is an ascription of a propensity, it will itself also have a certain degree of epistemic probability. Since PS is supported by our observations, and seems to be a reasonable thing to believe on the basis of those observations, PS will presumably have some degree of epistemic probability.

Now, let us suppose we continue to toss the dice, and find *that* "six" continues to come up about half the time. How are we to explain this? It is, as far as I can see, reasonable to explain it by saying the dice is weighted, or by saying that there is a propensity for "six" to come up that is about a half. PS can surely play an explanatory role. And the fact that PS *has* an epistemic probability does not prevent it from being able to play an explanatory role. On contrary – although PS would still be the type of statement that was capable of playing an explanatory role even if it had an epistemic probability of zero, it is surely rational for us to assert or believe PS as an explanation of the subsequent behaviour of the dice only if it does have some reasonably high epistemic probability.

Let us now consider again proposition: "(4) The more independent a theory is from the data the more likely it is that the data would conform to the theory if the data were to be obtained from other, including future, locations". The evidence for the truth of this proposition (4), it has been argued, is *a priori*. That is, (4) has some degree of epistemic probability. The degree of epistemic probability of (4) depends on the merit, or otherwise, of the *a priori* argument for it. We will not here go in to the merits or otherwise of the argument: I have done so elsewhere. But let us, for the sake of the argument, accept that (4) has some epistemic probability. Now, let us assume that T is some theory with a high degree of independence from the data. So, we will accept:

T has a high degree of independence from the data._________(2)

It follows from (2), and the claim that (4) has a good level of epistemic probability, that:

It is epistemically probable that it is likely that future data will conform to T. _____(3)

But now, (3) makes it *reasonable to assert*:

shows that more independent theories have a higher probability of conforming to the data than theories with a lower degree of independence. The argument for the future predictive success of highly independent theories, being purely probabilistic, does not in any way rely on dubious inference to the best explanation.

This enables us to construct a purely probabilistic argument for at least *some* cases scientific realism. This argument proceeds as follows. First, data is obtained concerning the behaviour of observable objects. We then try to discern some pattern or regularities in the behaviour of these observable objects. Given the underdetermination of theory by actual data, there will be a number of possible patterns we could attribute to the observable objects. We select the pattern that exhibits the highest degree of independence from the data. Let us call this pattern P. It follows from the argument sketched in this chapter (and defended in more detail elsewhere) that we have probabilistic reason to say that that the observable objects would conform to this patter P if we had obtained our data from other locations. And so we have probabilistic grounds for asserting:

Observable objects conform to pattern P.________(1)

But now, if observable objects conform to P, an Eddington inference entitles us to draw the conclusion:

There exist unobservable objects that conform to pattern P.___(2)

Moreover, the inference from (1) to (2) is also – it has been argued – purely probabilistic. And so we therefore have a purely probabilistic route from the data to a claim about the existence of and patterns or laws obeyed by *unobservable* objects. But, it should be observed, it does not automatically follow that these entities can be identified with the unobservable entities postulated by some explanatory theory T. On the view advocated here, it is the "No Coincidental Agreement" inference that enables us to carry out this identification. If the entities to which we are led by an Eddington inference would have the same observable effects as the observable phenomena explained by T, then the "No Coincidental Agreement" inference permits us to assert that that they are the same entities. Moreover, it was argued, use of his inference does not involve inference to the best explanation. Hence, we are able, on

---

It is likely future data will conform to T.______________(5)

Now, the sense of "likely" that appears in (5) is not epistemic probability but propensity. I have elsewhere argued that the more independent a theory is from the data, the more (epistemically) likely it is that there exists a propensity for the data to conform to the theory. So, at least if the argument for the preferability of independent theories is a good one, we may take (5) to be equivalent to:

There exists a propensity for the data to conform to T.________(6)

It is (6) which, on the view I have defended elsewhere, *does the explaining*. More specifically, it has been argued that it explains the novel predictive success of science. No claim is made that epistemic probability explains anything. All that epistemic probability does is justify our acceptance of claims such as (5) and (6). But the only type of probability that is asserted to actually explain anything is propensity.

this account, to arrive at Scientific Realist claims without inference to the best explanation.

6.7 Applying the Independence of Theory from Data to Actual Science

The independence of a theory T from data was earlier defined as the ratio of the number components of data explained by T to the number of dependent explanatory components of T. While we have discussed this notion at an intuitive level, we have as yet to offer definitions of "components of data" and "dependent explanatory components of theory". The aim of this section is go some direction in giving an account of these notions. It must be stressed that the following account is very broad-brush and incomplete. More detailed accounts of these notions have been given elsewhere.[18]

Let us begin by reviewing some aspects of our argument for Premise 1: "The more independent a theory is from the data, the less likely it is that it should merely be due to the blindly chosen location of the data that the data should exemplify a theory with that degree of independence." We argued for the truth of this premise by arguing that what we intuitively regard as single patterns are *a priori* likely to be rare in data. We explain some body of data by identifying patterns within it. Our strategy was to argue that bodies of data consisting of a small number of simple patterns were rarer than those consisting of many parts or components exemplifying different patterns. What is explained is some part or segment of the data that exemplifies a single pattern, and we explain it by showing how it does exemplify that pattern. This account of explanation is consistent with standard accounts of explanation in science. On, for example, the deductive-nomological view of explanation, we explain an individual event by showing how it can be subsumed under a general law. So, in applying the notion of independence to scientific theories, it is natural to take as a single explanatory unit a statement of a law. These will typically be universal generalisations. That is, we take a Dependent Explanatory Component (or "DEC") to be a single law.

What is to count, for our purposes, as a single component of data? At first, it might be suggested that we ought to take as a single component of data the record of an *individual* observation. But it is not too difficult to see that this would not be a good way of doing things. Consider, for example, the law-like generalisation: "All samples of water (at one atmosphere of pressure) boil at 100 °C". This plainly has very many positive instances: the number of observed positive instances of it plausibly runs in the millions. If we were to define a single component of data as an individual observation, the degree of independence of "All samples of water boil at 100 °C" (that is, the ratio of components of data to DECs) would be close to the

[18] See *Science and the Theory of Rationality* and *Explaining Science's Success*.

maximum possible value. But this does not seem quite right. In science, theories achieve a particularly impressive degree of independence when they show how what appear to be a number of distinct regularities at the observational level can be seen as instances of a significantly smaller number of regularities at the theoretical level. For example, the several regularities implicit in the gas law PV = *n*RT and in phenomena such as Brownian motion can be explained by saying gases are made of tiny masses elastically colliding and moving around according to Newton's laws of motion. The theory that gases are made of objects moving around according to Newton's laws achieves independence from the data by showing how wide range of regularities can be explained by a fewer. But if independence is achieved by showing how a larger number of regularities can be explained in terms of a smaller, it is natural to suggest that a single component of data is to be taken as an empirical *regularity* rather than as a single observed event.

This suggestion is supported by the fact that we regard a theory as having great explanatory power to the extent that it explains a large number of regularities, or types of phenomena, rather than a large number of individual events. Suppose we make a fairly small number of observations of the melting point of silver. For definiteness, let us assume we have observed five samples of silver. We are convinced the samples of silver we have observed are chemically pure. Suppose all the samples we have observed have melted at 1, 100 °C. Then we would be fairly confident that *all* samples of silver, including the (possibly infinitely many) unobserved samples, would all melt at 1, 100 °C. And so we are, surely, confident that the generalisation "All silver melts at 1, 100 °C" not only can explain the five actually observed meltings of silver, but would be able to explain many, possibly infinitely many, other similar events.

Now, let us assume we have also observed the temperature at which the metal tin melts. Let us assume we have found it to melt at 900 °C. Assume that we have observed 50 cases of silver melting. Again, we will suppose we are satisfied that the samples are chemically pure. We would – as with the case of silver – be confident that all samples of tin melted at 900 °C.

So, in summary, we are assuming in our examples that we have observed five cases of silver melting and fifty cases of tin melting. Would we, under these circumstances, be inclined to draw the conclusion that "All samples of silver melt at 1, 100 °C" had ten times the explanatory power of "All samples of tin melt at 900 °C"? I think it is clear we would not. We would, I think, say that both generalisations had (more or less) the same explanatory power. And the reason for this is intuitively quite clear. Not only does each generalisation explain some limited number of actual events (five in one case, fifty in the other) we are also confident that each generalisation would correctly explain indefinitely many other events. Each generalisation, we are confident, has the power to explain one empirical *regularity*. Since each generalisation can explain one empirical regularity, we are inclined to say they have (more or less) the same explanatory power.

These considerations suggest that it is natural to take as a single component of data a single *empirical regularity*.[19]

So, in summary, a single dependent explanatory component of theory is to be taken as a single, law-like generalisation, while a component of data is to be taken as a single empirical regularity. The degree of independence of a theory is therefore the ratio of the number of empirical regularities explained by it to the number of law-like generalisations employed by it.[20]

We can now describe some general features of the notion of independence. If a theory explains a single empirical regularity with one explanatory, law-like generalisation, it will have a degree of independence of one. Intuitively, such a theory will have not achieved any degree of independence from the data: it will simply a description of the data.[21] A theory starts to be a good one only when its degree of independence rises above one, that is, when it postulates fewer lawlike regularities than exist in the empirical data.[22]

On the approach advocated here, we establish realism with respect to some theory T by identifying the laws obeyed by some class of entities at the observational level and then using an Eddington inference to extend this to unobservable entities. But, of course, if such an approach is to lead to realism with respect to actual scientific theories, the laws which, on the present account we identify as those governing the behaviour of observables must also be those which has a matter of fact scientists have selected. And so the question arises: Do scientists as a matter of fact prefer theories that exhibit a high degree of independence from the data?

---

[19] This is discussed in more detail in my *Explaining Science's Success*, pp. 78–86.

[20] As noted in the main text, this is only a very broad-brush account. More detailed accounts are given in Wright (1991, 2013). In Wright (1991) an account is given which includes both causal claims and existential claims as explanatory components of theories.Mario Alai, *op cit*, objects that the account of independence given in Wright (2013) only applies to laws and not to theories. The main focus of his objection seems to be that the account of independence only applies to generalisations, and not to existential and causal claims. While this is true of Wright (2013), these matters were treated in Wright (1991). (They were not included in the later work for reasons of space.)

[21] This suggests that there might be something to be said for offering a slightly different definition of independence. It may for certain purposes be more appropriate to define independence as follows:

$$\text{Independence of T} = \frac{\text{Number of CODS explained by T}}{\text{Number of DECs of T}} - 1$$

On this definition, a "theory" that is nothing more than a description of the regularities to be found in the data will have a degree of independence of zero. A theory that is even more complex than the data it purports to explain will have a negative degree of independence. A theory only starts to be a good one once its degree of independence has some positive, greater than zero value.

[22] A more detailed account of the individuation of dependent explanatory components of theory (DECs) and components of data (CODs) can be found in the author's Science and the Theory of Rationality (Avebury, 1991) and Explaining Science's Success: Understanding How Scientific Knowledge Works (Acumen Publishing, 2013)

It has elsewhere been argued that there is considerable evidence indicating that scientists do prefer highly independent theories. The author has argued that Newton's Rules of Reasoning in Natural Philosophy can be seen as strategies for maximising the independence of theory from data. His arguments for his three laws of motion and his argument for Universal Gravitation can all be explained if we see Newton's as always making inferences to those theories that are most independent of the data they seek to explain.[23] The transition from phlogiston theory to Lavoisier's oxidation theory of combustion can be explained if we see the scientific community as exhibiting a preference of explanations that maximise independence from data.[24] Einstein's arguments for the Special Theory of Relativity can be seen as a series of inferences in which the most independent theory is accepted.[25] The arguments of Gregor Mendel from observations of peas to his theory of genetics can also be seen in this way.[26] So: the notion of independence can account for a number of key episodes in the history of science.

It has also been argued that the preference that scientists have for many of the properties generally regarded as good-making properties of theories can be seen as naturally following from a desire on the part of scientists to maximise independence. More specifically, it has been argued that the preference scientists exhibit for theories with simplicity, falsifiability, elegance, unity, great empirical content, symmetry, fruitfulness, a high degree of coherence with other theories, and for theories that are in agreement with metaphysical beliefs can all be seen as following from an underlying desire to maximise the independence of theory from data.[27]

6.8 Concluding Remarks

The aim of this chapter has been to argue that we can have probabilistic reason to prefer one way of discerning a pattern or regularity in the data to another. It has been argued (largely drawing upon earlier work) that the notion of independence of theory from data gives us this probabilistic reason. It has also been noted that this notion successfully applies to a number of examples from the history of science. Perhaps most importantly, the notion of independence of theory from data successfully applies to Newton's laws of motion, in the sense that those laws are more independent of the data than other possible ways of accounting for the observations of objects moving around in our terrestrial environment. This will be important in the next chapter, as the assumption that we have probabilistic reason for Newton's laws an essential part of the argument for the thesis that we can have good, probabilistic reason for the existence of molecules.

---

[23] See *Explaining Science's Success,* Chap. 6.

[24] This is argued for in *Science and the Theory of Rationality* (Avebury, 1991), Chap. 6.

[25] *Explaining Science's Success*, Chap. 7.

[26] *Explaining Science's Success*, Chap. 8.

[27] *Science and the Theory of Rationality*, Chap. 5.

Chapter 7
Eddington Inferences in Science – 1: Atoms and Molecules

7.1 Summary of Conclusions So Far

Let us begin by reviewing our main results to this point. In Chap. 2 it was argued that it is possible to give a probabilistic justification of induction. In Chap. 5 it was argued that a probabilistic argument, resembling an inductive argument, can also be given for the reality of some unobservable entities. The type of inference used in arguments of this sort was called an "Eddington-inference". But, it was also argued that if Eddington-inferences were to furnish us with good reason to believe in the unobservable entities postulated by specific scientific theories, they needed to be supplemented with a means of determining that one theory *about the behaviour of observables* was more likely to be true than another. It was argued that the notion of the independence of theory from data was able to do this. The viability of that notion for the task at hand was defended in Chap. 6.

It is worth observing that, on the account advocated here, the key concepts used: induction, Eddington-inferences, and the independence of theory from data, are all justified in a similar way. The arguments used in each case have the same structure: It is first observed that, given that the location of our data was blindly chosen, it is unlikely to be merely due to chance that our data should have some property S. The inference is made that it is probably not due to chance that it has S. Finally, from the fact that it is probably not merely due to chance that our data has S, it is concluded that a particular assertion S* has an increased likelihood of being true.

All the forms of inference used here are justified in the way described above. In the case of induction, we begin with the assertion that, given that the location of our observations was blindly chosen, it is unlikely to be merely due to chance that all observed crows are black. The inference is made that it is probably not due to chance that all observed crows are black, and hence the chances that are crows are black is thereby increased. In the case of the Eddington inference, we began by noting that, since the size of the holes in our fish trap was blindly chosen, it is highly unlikely that the size of the smallest fish in the sea would have happened to have coincided

J. Wright, *An Epistemic Foundation for Scientific Realism*, Synthese Library 402, https://doi.org/10.1007/978-3-030-02218-1_7

with the size of the holes in our trap. The conclusion was drawn that there probably are smaller fish in the sea. In the case of the inference to the theory most independent of the data, we began with the observation that, given the blindly chosen location of the data, it is unlikely to be merely due to chance that our data should conform to a theory with such a high degree of independence as our theory T. The conclusion is drawn that the data probably have a propensity to conform to T, and therefore that data from locations different from the blindly chosen actual location will probably conform to T.

Let us now put together the results of Chaps. 2, 5, and 6 to state the conditions under which, according to our view, an inference to a scientific realist claim is justified. First some empirical data is obtained. We then look around for patterns or regularities in the data D. It has been argued that the more independent a theory that can account for the data, the more likely it is that there is a propensity for the data to conform to the theory. So, we have found some theory T highly independent of the data D then, it has been argued, we have probabilistic reason to believe the data will continue to conform to T. It has also been argued that scientists do, in fact, tend to prefer theories with a high degree of independence from the data. But the independence of T does not give us any reason to believe anything about the existence of, for example, entities too small to see. It is the role of the Eddington-inference to do this. Our observations, together with an inference to the most independent theory that can explain the data, gives us reason to suppose that entities large enough to be detected by our apparatus obey T. An Eddington inference leads us to the conclusion that there are probably also entities not detectable by our apparatus that obey T.

On the view advocated here, the "No coincidental agreement" inference entitles us to assert that the entities to which we are led by an Eddington-inference are the same as those postulated by some explanatory theory T*. If the entities to which we are led by the Eddington inference have the same observable effects as would the entities postulated by T* if they existed, then the "No coincidental agreement" inference tells us we are justified in asserting that that they two classes of entities are the same. And since we do have (probabilistic) reason to believe in the existence of the entities to which we are led by the Eddington-inference, we thereby also have reason to believe in the existence of the entities postulated by T*. Thus, on the account offered here, realism with respect to T* is justified.

The aim of this chapter is to apply this approach to a case from the history of science. The case to be examined concerns the existence of atoms and molecules.

7.2 Maxwell's Arguments, Newton's Laws and the Gas Laws[1]

In this section we examine the arguments for atoms given by James Clerk Maxwell. It will be argued that the arguments Maxwell uses have the same structure as the way of arguing for molecules sketched in Chap. 5.

[1] In this section I will only be discussing the arguments of Maxwell for the existence of molecules from the behaviour of gases. I will not discuss the work of Ludwig Boltzmann, even

In his "Molecules"[2] Maxwell begins by noting that any piece of matter – his example is a drop of water – is divisible. He asserts that the process of dividing the drop of water could be carried out until "the separate portions of water are so small we can no longer see them." He then says "…we have no doubt the sub-division might be carried out further, if our senses were more acute and our instruments more delicate".[3] Plausibly, Maxwell would here seem to be making what we are calling an Eddington inference. Is it the case that there are portions of water smaller than those we can see or manipulate that nonetheless still have the property of physical divisibility? To suppose otherwise would be to suppose that the actual limits of the "acuteness of our senses and the delicacy of our instruments" would happen to coincide with the limit of physical divisibility, and Maxwell says we have "no doubt" this is not the case. And so, he accepts there exist portions of water, or matter of any kind, too small to see, that are also subject to physical divisibility. Maxwell, that is, accepts an Eddington inference to the existence of unobservables.

Maxwell explains that his concern is not with the existence of atoms in the classical sense of portions of matter not susceptible of further physical division, but rather with *molecules* of chemical substances, which he defines as the smallest part of any chemical substance having the same chemical properties as that substance.[4]

The next step in Maxwell's argument is to note that, *according to the theory he advocates*, the molecules of gases and liquids are in motion. His reason for this would seem to be because the hypothesis that they are in motion provides the best

---

though he also advocated and developed atomic theory as the explanation for the behaviour of gases.

As far as I can tell, Boltzmann nowhere employs what we are here calling Eddington inferences. This might be thought to be something of an embarrassment for the view advocated here. If Boltzmann believed atoms were real but did not use Eddington inferences, might this not suggest there does exist some good, rationally persuasive argument for the existence of atoms that does not use Eddington inferences?

However, it seems to me that the case of Boltzmann actually does not constitute an embarrassment for the view advocated here. It is far from clear that Boltzmann adopted a "fully realistic" view of atoms. To be sure, he was an enthusiastic advocate of atomic *theory*, but the *interpretation* he gave of that theory is not entirely clear, and was *not obviously* realistic. According to John Blackmore, for example, Boltzmann at one point understood atoms in terms derived from J. S. Mill as "permanent possibilities of sensation". (See J. Blackmore *Ludwig Boltzmann*: *His Later Life and Philosophy: Book Two* (Kluwer Academic Publishers, 1995), p.73.) More generally, according to Blackmore, Boltzmann seems to have had general ontological views closely related to the positivism of Ernst Mach. Blackmore also states that at one stage in his career Boltzmann viewed atomic theory as something that could be viewed in a number of ways: realists could see atoms as literally existing, while opponents of realism could still use atomic theory as a useful model. The general picture that emerges from Blackmore's work is that Boltzmann was clearly *not* what we would nowadays regard as a committed realist. Boltzmann would appear to have accepted the broadly positivist view that seemed to be prevalent amongst German theorists at the time. See Blackmore, *op cit*, pp.65–77.

[2] See Maxwell "Molecules" reprinted in *Maxwell on Molecules and Gases* edited by Elizabeth Garber, Stephen G. Brush and C. W. Francis Everett, (MIT Press, 1986), pp.137–155.

[3] See Maxwell "Molecules" in Garber, E et al., *op cit*, p.139.

[4] Maxwell, *op cit*, p.140.

explanation of the observed facts of diffusion.[5] So, here Maxwell would seem to be relying on IBE.[6] However, it will be argued that this use of IBE does not mean that his argument for molecules relies on IBE in any way inconsistent with the view developed in Chap. 5. Let us briefly review the account we developed there. Suppose we have a hypothesis H about the entities out of which a gas is composed. Then, on the view defended, we are entitled to accept what H has to say about those entities if the causal powers H attributes to those entities are the same causal powers as those to which we would be led by an Eddington inference, together with non-*ad hoc* hypotheses about the nature of such entities, if those entities were what gases were made of.

Now, the hypothesis that the molecules of a gas are in motion – whether or not it was for Maxwell rationally justified – is clearly not *ad hoc*. The phenomenon of diffusion would certainly *suggest* that the parts out of which gases are made are in motion, as would the theory that heat is the motion of the tiny components of matter.

It is worth reminding ourselves that, on the view advocated here, no assumption is made that the lack of *ad-hocness* of a theory indicates likely truth. Its role is merely to ensure the unlikelihood of agreement between our theory and an Eddington inference. If the entities to which we are led by an Eddington inference have the same causal powers as those postulated by our explaining theory, together with non-*ad hoc* hypotheses, then it is highly unlikely that such agreement should merely be due to chance. And it was argued in Chap. 5 that this gives us reason to say that the entities postulated by our explaining theory probably do exist.

So, in summary, while it is true Maxwell would appear to rely on inference to the best explanation to arrive at the claim that the molecules of a gas are in motion, this need not vitiate the approach advocated here. For us, it is sufficient that this hypothesis not be *ad hoc*.

Maxwell then says that if gases were composed of portions of matter too small to see that were also in motion, then they would have certain properties. Amongst these properties would be a propensity to conform to Newton's laws of motion. Although in "Molecules" Maxwell does not explicitly derive the conclusion that these portions of matter would conform to Newton's laws of motion via an Eddington

[5] Maxwell, *op cit*, p.141.

[6] Maxwell's wording suggests that he is relying on IBE at this point, but it seems there is a plausible route from the phenomenon of diffusion to the conclusion that the particles are in motion that does not use IBE. Maxwell says that when he opens the stopper on a flask containing a strongly smelling substance, "fairly quickly" it becomes possible to smell the substance from some distance away. This certainly would seem to indicate that once the stopper on the flask had been removed, the particles comprising the substance contained therein were in motion: they *moved* from the mouth of the flask to distant points in the lecture theatre. But now, unless the act of removing the stopper from the flask imparted motion to the particles, the law of the conservation of energy assures us the particles must surely have also been in motion prior to the stopper being removed. This argument assumes the law of conservation of energy, but this would appear to be something that can be confirmed without having to appeal to any problematic form of IBE. And so we seem to have a possible route to the conclusion that the particles were in motion, prior to the removal of the stopper, that does not rely on problematic IBE.

inference, this inference would seem to be implicit in what he says. Let us look closely at exactly what he says:

> Take any portion of matter, say a drop of water, and observe its properties. Like every other portion of matter we have seen, it is divisible. Divide it in two, each portion appears to retain all the properties of the original drop….The parts are similar to the whole in every respect except in absolute size.
>
> Now go on repeating the process till the separate portions of water are so small we can no longer perceive or handle them. Still, we have no doubt that the subdivision might be carried further, if our senses were more acute and our instruments more delicate."[7]

In the first paragraph cited above, Maxwell says that each new portion of the water has all of the properties of the original. Since the original drop of water had mass, it follows from what Maxwell says that each new portion will also have mass. But now, there was available to Maxwell reason to believe that all masses obey Newton's laws.[8] So, it follows that, on Maxwell's view, there is reason to believe each new portion of water will obey Newton's laws. Maxwell goes on to say that "we have no doubt the subdivision might be carried further". As noted above, in making this move Maxwell is plausibly making an Eddington inference: if things too small to see were not divisible the limit of divisibility would have happened to have coincided with the limit of perceptibility, which seems *a priori* unlikely. And so it follows, for Maxwell, that there probably exist portions of matter too small to see that have the same properties of matter that we can see. So, these portions of matter will have mass and conform to Newton's laws of motion. Moreover, a restricted Eddington inference gives us reason to say that the portions of matter as small as molecules (however small that may be) have mass and obey Newton's laws.

Maxwell then proceeds to derive some consequences from the supposition that gases consist of many of these small portions of matter in motion. The first consequence he derives concerns the relation between density and pressure. Clearly, the greater the amount, or mass, of gas that is put in to a vessel of given volume at constant pressure, the greater will be the density. But also, the greater the amount of the gas put in to the volume, the larger will be the number of the individual portions of matter that will, in any unit of time, strike against the walls of the vessel containing that volume of gas. Provided other properties of the gas – including its temperature – remain constant, doubling the quantity or mass of gas will result in doubling the density, and also doubling the number of particles that, in any given time, will strike against the walls, tripling the quantity will triple the number of strikings, and so on. And so Maxwell has derived what he refers to as "Boyle's law" that the density of a gas in a given volume, at a given temperature, is proportional to the pressure.[9]

---

[7] See Maxwell, *op cit*, p.139.

[8] The present author has also argued that the reasons available to Maxwell gave him *probabilistic* reason to prefer the hypothesis that all masses conformed to Newton's laws. (*Explaining Science's Success*, ch 6.)

[9] Maxwell, *op cit*, p.142. Boyle's Law is perhaps more commonly expressed as the law that, for any given mass of gas, pressure is inversely proportional to volume. However, it is plain that the law

Maxwell then moves on to a discussion of Charles' Law relating the pressure of a gas to its temperature. It is worth examining Maxwell's brief discussion of this matter closely, at it proves to conform to the model given here in Chap. 5. The initial premise of his argument is that the particles of matter of which a gas is composed are in motion. He then asks us to consider what would happen if the rapidity with which the particles are moving is increased, more specifically, doubled. First, he notes that each particle would strike against the walls of the vessel with twice the speed. But, since each particle is now moving at twice the speed, the *time* between successive impacts with the wall be halved, and so the *frequency* with which a particle strikes the wall will also be doubled. The overall effect of doubling the speed with which the particles are moving is to increase the force they exert fourfold. Maxwell draws the more general conclusion that if the velocity of each individual molecule in the gas is increased by some factor N, the resulting pressure of the gas will be increased by *the square of* N.[10]

In summary, there is an Eddington inference, together with a non *ad-hoc* hypothesis, that leads us to the conclusion that gases are composed of particles of matter too small to see that are in motion. Another Eddington inference leads us to say these particles would be subject to Newton's laws of motion. So, purely probabilistic inferences, together with a non *ad-hoc* hypothesis, tell us there exist in gases tiny particles subject to Newton's laws. Maxwell has considered what would happen if the rapidity of the motions of these particles were increased. He has argued that the force that these particles would exert on the walls of the containing vessel would vary as *the square* of the velocity of particles.

Maxwell then discusses his favoured theory of the nature of heat. According to the theory of heat to which he subscribes, the heat of an object is actually the continual motion of tiny particles of matter out of which the heated object is composed. To increase the temperature of an object is therefore to increase the rapidity with which these particles are moving. And so, he argues, this theory of heat provides a simple and natural explanation of Charles' Law that, if volume is kept constant, the pressure of a gas is proportional to its temperature.[11]

There is one more fact about heat that needs to be noted. Other experimental work had already revealed to Maxwell that the temperature of a gas is equivalent to *energy*.[12] Moreover, as we have just noted, for Maxwell, heat is the motion of tiny particles of matter making up the heated object. But now, the kinetic energy of an object in motion varies as *the square of* the velocity of that object. So, if heat is the motion of tiny particles, and heat is equivalent to energy, the temperature of an object will be proportional to *the square of* the velocity of those tiny particles.

Maxwell is here referring to as "Boyle's Law" entails that pressure will be inversely proportional to volume for a given mass of gas.

[10] See Maxwell, *loc cit.*

[11] Maxwell, *loc cit*.

[12] See J. P. Joule "On the Existence of an Equivalent Relation between Heat and the Ordinary Forms of Mechanical Power" in *Philosophical Magazine*, 3 **27** (1845), pp.205–207.

These results lead naturally to the conclusion that the entities to which we are led by the above Eddington inferences are one and the same entities as those postulated by the theory of heat favoured by Maxwell. Moreover, this identification conforms to the account given in Chap. 5. We have been led by Eddington inferences to postulate the existence of tiny particles of matter making up gases, and that these tiny particles possess certain properties. One of these properties is that they are in motion, and the square of the speed with which they are moving would be proportional to any force exerted on the walls of a vessel containing them.

According to the theory of heat Maxwell favours, the tiny particles out of which a heated body is composed will have just those properties: that is, the force they exert against the walls of any containing vessel would be proportional to the square of the velocity of those tiny moving particles. This follows straightforwardly from results already noted. The temperature of a gas is a measure of its energy. If heat is, as Maxwell's theory asserts, a form of motion, it is a form of kinetic energy. The kinetic energy of an object or a collection of objects is proportional to the square of the velocity with which it its constituent parts are moving. So, according to the theory of heat to which Maxwell subscribes, the temperature of an object would be proportional to the square of the velocity of the tiny particles making up a heated object. But now, Charles' Law tells us temperature is proportional to pressure. So, on the theory of heat favoured by Maxwell, the pressure a heated gas will exert on the walls of any containing vessel will be proportional to the square of the velocity of the tiny particles the motion of which is heat.

So, in summary, Eddington inferences lead us to the existence of certain entities (particles in motion) that have the same properties as those postulated by Maxwell's theory of heat. The square of the velocity of these particles is proportional to the force they exert on the walls of containing vessel, and this is just the property that would be possessed by the particles making up heated matter, if heat were the motion of such particles. By the no coincidental agreement inference we may therefore assert that the motion of the particles to which we are led by an Eddington inference is one and the same thing as that postulated by Maxwell's theory of heat.

The key idea here may perhaps be expressed as follows: Eddington inferences lead us to say entities with certain properties exist. (The properties are being too small to see, in motion, and such that the square of their velocity is proportional to the pressure they exert on containing walls.) Since Eddington inferences are probabilistic, we have good grounds for saying these entities exist "anyway", independently of any explaining theory. But now, Maxwell's favoured theory of certain actually observed phenomena associated with heat postulates the existence of entities with these very same properties. The no-coincidental agreement principle asserts that we are entitled to say that the entities postulated by Maxwell's theory of heat are one and the same as those which our Eddington inferences assure us probably exist "anyway", independently of the phenomena explained by our theory.

To summarise the results of this section, Maxwell has given arguments for identifying the motions of particles too small to see with both the cause of pressure, and temperature. The account given conforms to the model advocated here.

7.3 Einstein and Brownian Motion

An important next step in making the case for the existence of atoms was Einstein's work on the interpretation of Brownian motion.[13]

The term "Brownian motion" refers to the scientist Robert Brown who, in 1827, observed that, viewed under a microscope, grains of pollen in water moved in an apparently random manner. Einstein explored the idea that Brownian movement of particles suspended in a heating liquid are just what we would expect if the theory that heat was the motion of tiny particles making up the heated liquid were correct. It is worth noting that Einstein did not assert that Brownian motion *was* due the motions of these particles, since – in his words – the "information regarding [Brownian motion] available to me (Einstein) is so lacking in precision I am unable to form any judgement on the matter".[14]

Early in his paper Einstein makes some brief remarks that can be interpreted as the claim that the kinetic-molecular theory of heat is supported by what we are here calling an Eddington inference. He writes: "According to the [molecular-kinetic theory of heat] a dissolved molecule is differentiated from a suspended body solely by its dimensions, and it is not apparent why a number of suspended particles should not produce the same osmotic pressure as the same number of molecules".[15] It is, admittedly, not entirely clear whether Einstein is here merely noting a consequence of the molecular-kinetic theory of heat, or saying that its plausibility is increased by an Eddington inference. But he says that, on the theory, a dissolved molecule is "differentiated from a [visible] suspended body solely by its dimensions". That is, according to the theory, those properties possessed by visible suspended bodies ought to also be possessed by those bodies in a liquid too small to see: whether they are visible or too small to see ought to make no difference to the properties we attribute to them. And this can naturally be seen as an Eddington inference from the observable to the unobservable. The conclusion, roughly, is that all bodies in a liquid – whether observable or unobservable – ought to behave in the same way and therefore have the same kinds of effects. And from this he draws a more specific conclusion: that a collection of visible suspended particles can be expected to produce the same osmotic pressure as the same number of molecules.

It seems clear that Einstein is adopting here a stance that is in keeping with the assumptions underlying Eddington inferences. This can be summarised as follows: the distinction between that which can be observed and that which cannot has no ontological or physical significance, objects that are (only just) large enough to see obey the same laws as those too small to see, and the observed behaviour of those things large enough to see entitles us to draw conclusions about those things that are not.

[13] See Albert Einstein "On the Movement of Small Particles Suspended in a Stationary Liquid Demanded by the Molecular Kinetic Theory of Heat" reprinted in *Investigations on the Theory of the Brownian Movement* edited by R. Furth (Dover Publications, 1956), pp.1–18

[14] Einstein, *op cit*, p.1.

[15] Einstein, *op cit*, p.3.

The primary aim of Einstein's paper was to examine what consequences would follow for the motion of a microscopically visible particle suspended in a heated liquid if the kinetic molecular theory of heat were assumed to be correct. We need not concern ourselves with the precise details of his investigations, but his main findings were as follows: Suppose O is a microscopic but visible object being struck, from all directions, by smaller objects moving "at random". Then, over some sufficiently small period of time Δt, the number of impacts will probably not be exactly the same from all directions, but will exert an overall force more in one direction rather than the others. This will cause O to be moved from its original position. Over the next short interval of time Δt^* the number of impacts on O will again probably exert more force in one direction, again causing O to move. But, the direction in which O will move during Δt^* will be random with respect to, and so probably different from, the direction it moved during Δt. Similar processes will occur in similar short periods of time. This will result in O taking a random path through the liquid, sometimes referred to as a "drunkard's walk".

Einstein derived a mathematical formula attributing certain properties to the path that would be taken by such a hypothetical body O. The formula showed how the extent of the movement of the body O over time would be determined by the temperature and viscosity of the fluid in which it was suspended, the size of the body itself, and Avogadro's Number.[16]

Einstein's result furnishes us with a possible way of testing the molecular kinetic theory of heat. Brownian motion is the agitation of tiny bodies suspended in a heated liquid. That the suspended bodies should become agitated in this way is broadly in keeping with the molecular kinetic theory of heat: it is "consistent with it", as that phrase is sometimes used. But in the absence of any *precise* description of how the suspended bodies would be expected to move if the molecular kinetic theory were true, Brownian motion cannot be claimed to provide strong confirmation of that theory. But it is just this *precise* description of how suspended bodies would be expected to move that has been provided by Einstein. If Brownian particles can be found to move in accordance with Einstein's formula relating movement to particle size, liquid temperature and viscosity and Avogadro's number, then the molecular kinetic theory of heat would appear to have successfully passed a test.

What Einstein has, quite clearly, provided is a possible way of *testing* the kinetic-molecular theory of heat. If suspended particles were found *not* to move in accor-

[16] The formula Einstein derived can be expressed as follows:

$$\mathrm{N} = \left(1/\mathrm{x}^2\right)\left(\mathrm{RT}/3\pi\eta r\right)\tau$$

Where N is Avogadro's Number, Spiltx^{2}Spigt is the mean square displacement of a particle over some unit of time τ, R is the gas constant, T is the temperature in degrees Kelvin, η is the viscosity of the liquid, and r is the radius of the particles. This expression of the formula comes from "Einstein, Perrin and Avogadro's Number – 1905 Revisited" by Ronald Newburgh, Joseph Peidle and Wolfgang Rueckner *American Journal of Physics*, **74**, 478 (2006).

dance with Einstein's formula, the kinetic molecular theory would appear to be in trouble. But suppose the particles were found to move in a way that accorded with Einstein's formula. Would this *confirm* the theory? On the face of it, to say that it would provide positive confirmation of the theory would seem to presuppose IBE. And, *prima facie* at least, Einstein appears to be saying that if Brownian motion were in fact found to possess the features described by his formula, the best explanation of this would be that given by the molecular kinetic theory of heat, and this would confirm the theory. That is, Einstein seems to be relying on IBE: a position in some tension with that advocated here. But it will be argued that looking at things more closely reveals that Einstein's approach is in fact in accordance with the approach advocated here.

In his derivation of the formula describing how microscopic but visible particles would move if the molecular kinetic theory of heat were true:

(i) Einstein assumes that the invisible molecules striking against the microscopic body obey the gas law $PV = nRT$, and Stokes' Law, relating the force required to move an object through a liquid of known viscosity.[17]
(ii) He also employs mathematical techniques from statistics.

In considering whether Einstein's conclusion can be reached via an Eddington inference, we need not concern ourselves with the purely mathematical/statistical part of his derivation. Our concern is only with ampliative inference. So, we need only concern ourselves with (i): his assumption that the particles would obey the gas law and Stokes' Law.

We saw in the previous section how Maxwell was able to derive the gas law on the model advocated in Chap. 5. Eddington inferences, together with non *ad-hoc* hypotheses, led us to say that there are particles too small to see and they move around according to Newton's laws of motion. From these assumptions Maxwell was able to derive Boyle's Law and Charles' Law. And these laws, together with Avogadro's Law, are sufficient to derive the general gas law Einstein uses in his argument.[18]

---

[17] Stokes' Law can be expressed as $F_d = 6\pi\mu RV$, where F_d is the force required to move a sphere of radius R at velocity V through a fluid of viscosity μ.

[18] Boyle's law states that, for an ideal gas at constant temperature, pressure is inversely proportional to volume, that is, $P = k_1(1/V)$, where P is pressure, k_1 is a constant, and V is volume. Charles' Law states that when pressure is held constant, volume is directly proportional to temperature, at constant pressure, $V = k_2 T$, where V is volume, k_2 is a constant and T is absolute temperature. Avogadro's Law states that equal volumes of all gases, at the same temperature and pressure, have the same number of molecules. This law can be stated as $K = V/n$, where K is a constant equal to RT/P, where R is the universal gas constant, T is absolute temperature and V is volume.

An objection might be raised at this point. Our overall aim is to show that there is a good, purely probabilistic, argument for the existence of molecules. The results established by Einstein are a step on the way to that conclusion. Einstein uses the general ideal gas law $PV = n\mathrm{RT}$, but this is derived from, among other things, Avogadro's law which makes reference to the number of *molecules* in a gas. It might be objected that if we have good grounds for saying that Avogadro's Law is true, then we must already have good grounds for saying that there are molecules (since Avogadro's Law makes reference to the number of molecules in a gas.) But if we already have

Stokes' Law introduces some further complications. Stokes' Law relates the force required to move a sphere of known size at a given velocity through a liquid of known viscosity. On the face of it, Einstein's use of this law might seem to present no difficulty: after all, properties such as the size of some spheres, their velocity, and the viscosity of a liquid can all be measured. We can verify at least *some* positive instances of Stokes' Law empirically. However, Stokes' Law had only been confirmed for bodies down to a certain size. So, its extension to other, smaller bodies might seem questionable. But, of course, given the role Einstein's results play on the approach advocated here, it is only required that Einstein's use of Stokes' Law not be *ad hoc*. And, clearly, Einstein's use is not *ad hoc*: on the view defended in this book the fact that Stokes' Law had been found to apply to certain observable bodies gives probabilistic support to the thesis it also applies to bodies too small to observe.

So, in summary, Eddington inferences (and non *ad-hoc* hypotheses), together with mathematical techniques of inference, are sufficient to derive Einstein's formula predicting how a suspended microscopic particle would move if the molecular kinetic theory of heat were correct. This of course raises the question: Do particles exhibiting Brownian motion actually move in accordance with Einstein's formula? If it were to be shown that they do, then the No-Coincidental-Agreement inference would justify us in concluding that particles exhibiting Brownian motion are in fact affected by tiny particles moving in the same way as that predicted by the molecular kinetic theory of heat, and therefore that theory of heat is correct. The procedure Einstein advocates is therefore in agreement with the one defended here.

7.4 The Experiments of Perrin

Jean Perrin was engaged in experimental work designed to determine the value of Avogadro's Number before he became aware of Einstein's work. The experiment for which he is best known was also performed before he became aware of Einstein's contribution. But, like Einstein's, it relied upon earlier work by Maxwell and others to test the molecular kinetic theory of heat.[19]

---

good grounds for saying molecules exist, is not the subsequent work of Perrin, for example, thereby rendered superfluous? And if we do not already have good reason for Avogadro's Law, is not the subsequent reasoning rendered unsound?

However, this objection fails. On the view advocated here, the "No-coincidental agreement" inference has a role in establishing the existence of unobservable entities. For this inference to work, we must be led via an Eddington inference together with non *ad-hoc* hypotheses to entities that have the same causal powers as those postulated by our explain theory. So, for our purposes, it is sufficient that Avogadro's Law not be *ad hoc*. And, it surely is not *ad hoc*: it provides a natural explanation of Gay-Lussac's Law that at any given temperature and pressure the ratio between the volumes of reacting gases and their product can be expressed in simple whole numbers. Gay-Lussac's law evidently could be established by enumerative induction.

[19] Perrin's main work summarizing the results of his findings is his *Atoms* (Constable and Company, London 1916).

On one view, Perrin's case for atoms relies on the agreement, or concordance, of a (large) number of methods for determining the value of Avogadro's Number. An early proponent of this view was Wesley Salmon.[20] Salmon's argument can be summarised as follows. Suppose that molecules did not exist. Then the probability of a wide range of different methods yielding (to a high degree of accuracy) the same value for Avogadro's Number would be astronomically low. But if molecules do exist, the probability may be quite high. But now, Perrin noted that as a matter of empirical fact, no less than thirteen independent methods had been found to yield, to a high degree of accuracy, the same value for Avogadro's Number.[21] We therefore have good reason to believe molecules exist.

However, as we have already argued, it is not entirely clear that this type of argument – and it is certainly at least *a part* of what Perrin says – gets us all the way to scientific realism about molecules considered *as tiny bits of matter*. As an argument for the existence of 6×10^{23} *somethings*, Salmon's reasoning seems very powerful indeed. But it would seem more argumentation is needed to show that those "somethings" are tiny bits of matter. To put the matter in terms of contrastive confirmation, Salmon's argument would seem to confirm that there are (approximately) 6×10^{23} "somethings" in a gramme of hydrogen, *rather than some quite different number*, but it is not so clear it confirms that those somethings are *tiny bits of matter*, *rather than entities of a different sort*. It would seem we need to do something more than just show Avogadro's number can be established by several independent methods if we are to show this.

It will be argued that Perrin's work *does* supply a way showing that atoms and molecules, considered as tiny bits of matter, do in fact exist. Very broadly, it will be argued that the different ways of determining Avogadro's Number, and some Eddington inferences, work together to furnish us with good reason for this.

Since the situation to be described is fairly complex, it is useful to first give a broad overview. It should be noted at the start that Perrin's *Atoms* is a long and detailed work, exhibiting many internal cross-references. The analysis to be offered below is only one part of Perrin's much more extended treatment.

A central part of Perrin's work concerned a particular law (to be described below), which we will refer to as *Laplace's Law of Atmospheres*. His investigations in to consequences of this law played a crucial role, but only a part of the role, in his overall argument for atoms. On the view to be defended, Perrin's argument for atoms can be presented as follows:

(a) A form of Laplace's law of atmospheres, together with some apparatus of Perrin's devising, made it possible to calculate a value for Avogadro's Number.

---

[20] See W. Salmon *Reality and Rationality* edited by P. Dowe and M. Salmon, (Oxford University Press, 2005), esp. pp.3–60.

[21] See Salmon, *op cit*. A summary of the results obtained by the use of the different methods is given in Perrin's *Atoms*, p.206. There Perrin lists thirteen different methods. The lowest value obtained (from a technique involving energy radiated in radioactivity) is 6.0×10^{23}. The highest is 7.5×10^{23}, from a technique involving "critical opalescence". The latter value would seem to be something of an outlier from the others. Perrin comments that the degree of agreement between the various methods is so remarkable "the real existence of the molecule is given a probability bordering on certainty." (p.207).

(b) Einstein's earlier work on Brownian motion provided another, independent way of determining Avogadro's Number.
(c) The two methods of (a) and (b) were in close agreement, confirming it as a method-independent fact that Avogadro's Number is approximately 6×10^{23}, rather than some quite different number.
(d) Laplace's Law of Atmospheres and Einstein's work on Brownian motion, together with the (now) confirmed value for Avogadro's Number, make it possible to construct Eddington inferences to the conclusion that atoms and molecules exist as tiny bits of matter in random motion, rather than as entities or powers of some different sort.

We will start by briefly looking at (a): How Perrin used a form of Laplace's law of atmospheres, together with an experimental set up of his own devising, to derive a value for Avogadro's Number.

Laplace's law of atmospheres relates the number of particles in a gas to height. It was derived by Laplace as an explanation of the familiar fact that, at high altitudes in mountainous regions, the air was known to be thinner. Very roughly, the law predicts that if a sufficiently large number of particles are in a gravitational field, then, at a greater height, the number of particles ought to be fewer, and at a lower height, it ought to be larger. Perrin derives a version of the law for particles suspended in a liquid.[22] It is as follows:

$$\frac{n^*}{n} = 1 - \frac{N}{RT}.m\left(1 - \frac{d}{D}\right)gh\underline{\text{Inner space}\left(\text{0xEF07}\right)}\left(\text{LA}\right)$$

Let A and B be two (arbitrarily chosen) points in the liquid, where the difference in height between A and B is *h*. Then *n** will be the number of the particles for a given volume at A while *n* will be the number of the particles for a given volume at B. So, the ratio *n*/n* will be the ratio of densities of particles in volumes separated by some height *h*. *N* is Avogadro's Number, *R* is the ideal gas constant, *T* is temperature in degrees Kelvin, *m* is the mass of the particles, *d* is the density of the liquid in which they are suspended, *D* is the density of the substance out of which the particles are composed,[23] *g* is the strength of the force of gravity and, as stated, *h* is the difference in height between the points at which the densities are given by *n* and *n**.

The derivation that Perrin employs makes no assumption about the *size* of the particles. It does not assume that the particles are "more or less" the size of molecules: they could be, for example, the size of ping-pong balls.[24] The derivation

[22] See Perrin, *op cit*, pp.90–94.

[23] Note that *n* and *n** refer to the "density of the particles" in the sense of the number of them to be found in a given volume of space. So, *n* will take a very high value if there are, for example, very many of the particles in a cubic centimetre. But D refers to the density of the substance out of which the particles are made. So, D will have a higher value if the particles are made out of, say, lead rather than aluminium.

[24] The derivation, and the formula, both contain reference to temperature in degrees Kelvin. But this need not be taken as assuming that *heat* is the kinetic energy of tiny particles. The term "T"

applies to any objects whatsoever, provided they obey Newton's Laws and conform to the ideal gas law $PV = nRT$.[25]

However, of course, these considerations would not entitle an investigator to assert that the law (LA) is actually true *of gases* unless it was *already* known that, in fact, gases *are* made up of tiny particles that obey Newton's laws. And the assumption that gases are made up of such particles seems to be pretty much equivalent to the hypothesis that molecules and atoms really do exist. So, it might seem to follow, any use of the law of atmospheres (LA) in attempting to establish the existence of atoms or molecules would simply beg the question.

However, it will be argued, this unwelcome conclusion need not necessarily follow. First, let us look again at our overall sketch of Perrin's argument, as given in (a)–(d). In Perrin's argument, the law of atmospheres (LA) is initially used, *in conjunction with* Einstein's findings on Brownian motion, to determine Avogadro's Number. In Perrin's argument, it is the fact these two methods yield the same value for Avogadro's Number that entitles us to conclude that Avogadro's Number *is, in fact* 6×10^{23}. Now, it is appropriate at this point to remind ourselves of a feature of the method of agreement or concordance between independent techniques as a way of establishing a result. If two techniques independently furnish us with the same result, then we may have good reason to accept the correctness of the result *even if*, prior to their use, we lacked reason to trust either method. (If two methods of determining the height of the mountain gave the same result, then we may have reason to accept what they say even if prior to using them neither method was seen as particularly trustworthy.) It is in this sense that the concordance between (LA) and Einstein's method does provide us with a way of determining Avogadro's Number. It gives us good reason to believe that there are (approximately) 6×10^{23} somethings in a mole of a substance. And it succeeds in doing this even if, at this stage, the reliability of neither Perrin's method nor the method of Einstein has been independently established. But, there is also something the method of concordance, considered in itself, *fails* to do: it fails to show those 6×10^{23} somethings are *molecules*, rather than something else.

Let us now look a little more closely at how (LA) might be used to give a value for Avogadro's Number. On the face of it, the method is perfectly straightforward. (LA) refers to "N" – Avogadro's Number – and nine other quantities. If we can determine the values of those other quantities, then we determine the value of "N", that is, of Avogadro's Number. But: *can we*, as a matter of practical fact, determine the values of the other quantities? Immediately, a difficulty presents itself. The quantities n and n^* refer to the number of the particles in a given volume, m refers to their mass and d to the density of the substance out of which they are composed. How are we to establish the values of these quantities if the particles are too small

that appears in the formula could be taken merely as a measure of the kinetic energy of the particles. So then LA might, for example, tell us how, at a given average kinetic energy, the density of a quantity of randomly moving ping pong balls diminishes with height.

[25] Again, assuming "T" to refer to kinetic energy.

to see? Perhaps Perrin's main experimental contribution was to work out a way around this difficulty.

As we have noted, the derivation of (LA), from the ideal gas law and ultimately from Newton's laws of motion, makes no reference to the *size* of the particles involved. And so, it ought theoretically to hold for particles of any size: whether they are, like atoms and molecules, too small to be seen, or whether they are larger. Perrin reasoned that, if the law of atmospheres were true, it ought to hold for tiny but microscopically observable particles suspended in a fluid. Assuming that the law did hold for such (microscopically) observable particles, it may then be possible to ascertain the values of the quantities and then use (LA) to determine Avogadro's Number.

Perrin made a suspension with tiny but microscopically visible particles of a resin called gamboge. Since they were microscopically visible, they could be counted. And so Perrin was able to establish the number of particles in a given volume at different heights in the suspension.[26] He was also able to measure the mass of the individual particles and the density of the gamboge resin. This made it possible to obtain a value for Avogadro's Number.

Note that, on the interpretation offered here, we are not yet entitled to assert that the obtained value for Avogadro's Number is correct. The method of calculation assumes that the particles of gamboge suspended in the fluid have distributed themselves in the way they have *because* of the collisions they are experiencing with other particles, too small to see. And the correctness of that assumption has yet to be established. But, given that the value of Avogadro's Number is to be determined by the agreement between independent methods, the fact that the reliability of this method has not yet been established does not matter.

As we have noted, another independent way of determining the value of Avogadro's Number was provided by Einstein's work. This method involved determining whether or not a particle exhibiting Brownian motion in fact possessed the features we would expect it to have if its motion were due to the impacts of tiny particles, in accordance with the kinetic molecular theory of heat.

Perrin conducted numerous detailed observations of the paths taken by particles of gamboge and mastic suspended in a variety of solutions.[27] He then performed a statistical analysis on the observed paths and used Einstein's proposed method to determine a value for Avogadro's Number. The obtained results were impressively close to the values derived by the use of (LA).[28]

The agreement between the two methods of determining Avogadro's Number provided good reason for believing that the value of (approximately) 6×10^{23} was correct. It gives us good reason to believe there are 6×10^{23} *somethings* in, say, a gramme of hydrogen, but it does not tell us just what those "somethings" are. For this, more argumentation is needed.

[26] See Perrin, *op cit*, pp.101–103.

[27] See Perrin, *op cit*, Chapter IV.

[28] See Perrin, *op cit*, p.123. Perrin remarks: "This remarkable agreement proves the rigorous accuracy of Einstein's formula and in a striking manner confirms the molecular theory."

Now that Avogadro's Number had been determined, however, the observations that Perrin made of particles of gamboge suspended in a liquid acquire new significance. Armed with Avogadro's Number, the observations of these particles now constitute confirmation that the particles *do in fact obey* the law of atmospheres (LA).

It might at first seem as though there is something "logically fishy" about this move, but closer inspection shows it to be justified. Initially – in step (a) – the law of atmospheres (LA) was used to calculate a value for Avogadro's Number. Now, we might be tempted to think that to use (LA), together with observations of particles suspended in a fluid, to derive a value for Avogadro's Number is surely to *assume* that that those particles do conform to (LA). And so, to then use the obtained value to *confirm* that the particles conform to (LA) may seem obviously circular. But this is not so. The initial *derivation* of a value for Avogadro's Number (from observing the particles of gamboge) was not, and need not be claimed to be, a *proof* or *demonstration* that Avogadro's Number had the value 6.022×10^{23}. To put the matter in logical terms, the derivation might constitute a *valid* argument, but its *soundness* was yet to be demonstrated. The assertion that Avogadro's Number was 6×10^{23} had the status, rather, of a risky hypothesis which was to be tested by deriving that number by another means – in this case the means suggested by the work of Einstein. When the two techniques were found to agree, the method of agreement of independent techniques permitted the inference to the conclusion that the obtained value for Avogadro's Number was correct. The confirmation that the suspended particles of gamboge obeyed the law of atmospheres came, perhaps slightly paradoxically, from the fact that the observations of Brownian motion turned out to yield the same value for Avogadro's Number. But there is no circularity in this.

So, in summary, once a value for Avogadro's Number had been confirmed, it became rational to assert that the suspended particles of gamboge were obeying the law of atmospheres.

But now we are confronted with the question: How does all this help to establish the existence of atoms and molecules? Here Eddington inferences are required. The two methods used to determine the value of Avogadro's Number – the method appealing to the law of atmospheres and the method developed by Einstein – each yield their own Eddington inferences. The inference yielded by the law of atmospheres is as follows:

Microscopic but just visible suspended particles in the liquid obey (LA).

Therefore: There are particles in the liquid smaller than the microscopically observable that obey (LA).

Perrin's result establishes that particles of gamboge conform to the law of atmospheres (LA). But particles of gamboge are only just large enough to be microscopically detectable. If particles only just large enough to be microscopically detectable conform to the law of atmospheres, Eddington inferences assure us both that there probably exist particles too small to see and that they too conform to the law of atmospheres.

The observations of the random paths taken by the suspended particles, together with the application of Einstein's method to those random paths, yields another Eddington inference:

Microscopic but just visible suspended particles of matter move in a random fashion in keeping with the properties predicted by Einstein.

There exist particles of matter too small to see that move in a random fashion in keeping with the properties predicted by Einstein.

It is this last Eddington inference that perhaps most directly supports the existence of molecules. In the section of *Atoms* titled "A Decisive Proof",[29] Perrin writes:

> The objective reality of the molecules therefore becomes hard to deny. At the same time, molecular movement has not been made visible. The Brownian movement [of the suspended particles of gamboge] is a faithful reflection of it, or, better, it is a molecular movement in itself, in the same sense that the infra-red is still light. From the point of view of agitation, there is no distinction between the [unobservable] nitrogen molecules and the visible molecules realised in the grains of the emulsion. Perrin, *Atoms*, p.105

The crucial parts of this passage are: "Brownian motion is…. molecular movement in itself, in the same sense that the infra-red is still light" and "there is no distinction between [unobservable] nitrogen molecules and the visible …grains of the emulsion". Here Perrin can be seen as appealing to an example, and to a type of reasoning, that we have already used in this book. In Chap. 5 we noted that an Eddington inference could be used to take us from the existence of visible light to the existence of infra-red and ultra-violet light. Perrin is here saying a similar type of relation exists between the Brownian motion exhibited by the suspended particles of gamboge and the motion of the (too-small-to-see) molecules. The particles of gamboge move in a way which, according to the equation developed by Einstein, has the same features as the way molecules would move if the kinetic theory of heat were correct. On Perrin's view, the motion of molecules is to Brownian motion as infra-red light is to visible light: they are instances of the same type of thing, with the sole difference that one is perceptible to us while the other is not.

It is time now to summarise all the above findings, and see how, together, they add up to an argument for existence of atoms. We have noted earlier how Maxwell provided an argument to the effect that heat was the motion of tiny particles moving around according to Newton's laws. One part of Maxwell's argument showed how it is possible to construct Eddington inferences, together with non-*ad hoc* hypotheses, to the conclusion that matter consisted of tiny particles that moved around according to Newton's laws and, in consequence, would obey the gas laws. We have also noted that if particles obey the gas laws then, under the influence of gravity, they will also obey the law of atmospheres. It is therefore possible to construct Eddington inferences (together with non-*ad hoc* hypotheses) to the conclusion that under the force of gravity particles of matter will obey the law of atmospheres. But

---

[29] See Perrin, *op cit*, pp.104–106.

now, Perrin's observations, together with an Eddington inference, lead us to the conclusion that there are tiny *unobservable* particles of matter in a liquid that do, *in fact*, obey the law of atmospheres. Einstein showed that if the kinetic theory were correct, then particles exhibiting Brownian motion ought to trace out paths with certain properties. Perrin showed that the path taken by particles of gamboge do exhibit these properties. And an Eddington inference takes us to the conclusion that there exist particles of matter too small to see that also move in this random manner. These two Eddington inferences lead us to the conclusion that there exist particles of matter too small to see that move around in the way that would be expected if the kinetic theory were true. And so, by the No-Coincidental Agreement Principle, we are led to the conclusion the entities to which we are led by the Eddington inferences, and the entities postulated by the kinetic theory of heat, are one and the same. And since we have probabilistic grounds for saying the entities to which are led by the Eddington inferences do in fact exist, we also have probabilistic grounds for saying that the entities postulated by the kinetic theory – that is, atoms and molecules – do, in fact, exist.

7.5 Defence of the Above Interpretation of Perrin as an Argument for Realism

In the previous section it was argued that Perrin's work does in fact provide us with a good argument for realism about molecules. But, of course, not all commentators have accepted that Perrin's argument is good. A frequently made accusation against Perrin is that his argument is circular, or somehow question-begging.

One author who has noted that Perrin's reasoning has at least the superficial appearance of circularity is Peter Achinstein.[30] Achinstein notes that that Perrin's argument uses as one of its premises Laplace's Law of Atmospheres, which – as we have already noted – makes a claim about molecules. The Law of Atmospheres appears to *assume* that molecules exist, which is the very thing Perrin is trying to establish. So, Perrin's argument has the appearance of circularity. But Achinstein argues that, on close analysis, Perrin's argument turns out to not be circular.

As Achinstein interprets Perrin, the latter's argument for the existence of molecules can be understood broadly as follows: It is an established fact that a number of different methods all yield the same value for Avogadro's Number. This number is (approximately) 6×10^{23}. Now, Achinstein says we may assert:

If molecules exist, then Perrin's experiment tells us that *there exist* in any mole of a substance 6×10^{23} *molecules*.______________(1)

Note that (1) does *not* assert that there exists in a mole 6×10^{23} molecules. It is rather merely a conditional that tells us that *if* molecules exist, then Perrin's method

[30] See P. Achinstein *The Book of Evidence*, (Oxford University Press, 2001), Chap. 12 "Evidence for Molecules: Jean Perrin and Molecular Reality", pp.243–265, esp. p.244.

tells us how many molecules there are in a mole. But now, as Salmon has argued, the concordance of a wide variety of different methods for determining the value of Avogadro's Number will make the hypothesis that molecules exist more likely than the hypothesis that they do not. We therefore have good reason to say molecules *do* exist. And so, by (1) together with *modus ponens*, we do now therefore have good reason to say that Perrin's method shows there to be 6×10^{23} molecules in a mole.

Achinstein argues that on such an interpretation of Perrin's reasoning, it is not circular. Perrin *does*, in one sense of the word, "use" the Law of Atmospheres in his method for determining Avogadro's Number. But his "use" of this law does not require him to assume it to be *true*, considered as a literally true description of how a system of molecules behaves in a gravitational field. Rather, Perrin accepts or uses the Law of Atmospheres in quite a different spirit: as telling us how molecules would behave in a gravitational field if, hypothetically, they did exist. And when Perrin records the result "6×10^{23}" after performing his experiment, he is not making the unconditional assertion that that is how many molecules there are in a mole, he is rather – as Achinstein interprets him – making something like the *conditional* assertion that if molecules do exist, *this* is how many molecules there would be in a mole. Construed in this way, according to Achinstein, Perrin's argument is not circular.

On the view presented here, the agreement between two (or more) independent methods can furnish us with good reason for the correctness of the result upon which they agree, even if prior to the fact of their agreement, we did not have good reason to believe in the reliability of either method. So, on the view presented here, we do not have to assume in advance that any one particular way of determining Avogadro's Number is reliable. In this respect, the view presented here is like that offered by Achinstein.

However, it has here been argued that the mere fact of agreement between independent methods is not by itself enough to establish the existence of molecules considered as tiny bits of matter obeying Newton's laws. And Achinstein is aware that mere agreement between the methods is not enough to establish scientific realism. He acknowledges that, at least on the face of it, a constructive empiricist interpretation of the observed empirical facts might also be possible. But Achinstein goes on to argue that, in fact, a constructive empiricist would *not* be able to give a satisfactory account of the agreement of the methods. More specifically, Achinstein says that if Perrin's atomic theory were to be given a constructive-empiricist interpretation, it would cease to furnish us with an *explanation* of the concordance of the different methods for determining Avogadro's Number. The fact that a wide range of different ways for calculating Avogadro's Number yield (to a high degree of accuracy) the same result is something that surely "cries out" for explanation. On the face of it, at least, one possible explanation is that the theory of molecules is true, in the sense of "true" defended by the scientific realist, and hence that molecules literally do exist and that the different methods are all different methods for counting the same, literally existing, entities. But, whether or not this explanation is the best – or indeed the only – explanation, it is plainly not available to the constructive empiricist. For the constructive-empiricist, we are not entitled to assert anything that goes beyond the claim the theory of molecules has empirical adequacy. But Achinstein argues that to say "The theory of molecules has empirical adequacy" would not seem to provide an *explanation* of the fact that the different methods agree: it is rather (among other things) merely an assertion of the unexplained fact they *do* agree.

We need not here consider whether Achinstein's claim against the constructive empiricist is correct, although in the opinion of the present author it is correct.[31] But it will be argued that *even if* Achinstein's claim about the explanatory ineffectuality of constructive empiricism in this context is correct, it does not quite follow that we are entitled to adopt a fully realistic view of molecules. Suppose we grant that the constructive empiricist cannot explain the facts of concordance, while the realist about molecules can. Does it follow that we are entitled to accept realism about molecules? Not necessarily. There might be other explanations for the concordance. And even if we say that the explanation appealing to molecules is *the best*, what entitles us to accept as true the best explanation? Further: even if it could somehow be shown that molecules provided the only explanation, we would still be confronted with the question: "What entitles us to assume that explanation is sufficiently good to warrant rational acceptance?" That is, we find ourselves once again confronted with the questions with which we were concerned in Chap. 4. I conclude that the account that Achinstein has given us does not get us all the way to realism about molecules.

On the view advocated here, we do not accept the theory of molecules because it is the best explanation of the phenomena. Instead, it is Eddington inferences, together with other probabilistically justifiable inferences, that justify the conclusion that there probably exist entities too small to see that possess the properties ascribed to molecules by the kinetic theory of heat. And the No-coincidental agreement inference entitles us to identify the entities to which we are led by the Eddington inference with those postulated by the kinetic theory. Eddington inferences enable us to complete the task of justifying realism about molecules in a way that the account offered by Achinstein does not.

According to Bas van Fraassen, defenders of scientific realism have misconstrued the historical, scientific and philosophical significance of Perrin's work.[32] For van Fraassen, Perrin's work is not, and neither was it ever intended to be, an argument for *realism* about molecules. Rather, its significance was quite different. For van Fraassen, Perrin's work merely supplies an *empirical grounding* for the theory of molecules.

In considering van Fraassen's position, it is useful to begin by briefly outlining what he says about the notion of empirical grounding. Under what circumstances is a scientific theory empirically grounded? A first suggestion might be:

> A theory is empirically grounded if there exists an experimental procedure for determining the magnitude or value of each theoretical property postulated by the theory.

---

[31] See for example my *Realism and Explanatorily Priority* (Kluwer Academic Publishers, 1997) and "The Explanatory Role of Realism" in *Philosophia*, v 29 (2002), pp.35–56.

[32] See B. van Fraassen "The Perils of Perrin, in the hands of philosophers", *Philosophical Studies*, v 143 (2009), pp.5–24.

So, the property of force, for example, would count as empirically grounded if there existed some procedure for determining the magnitude or value of the force acting on an object.

Van Fraassen argues that this will not quite do. He illustrates his point by reference to certain devices that were held by Newtonians to furnish us with a way of measuring force, but which seemed to *presuppose* the correctness of Newton's laws.[33] For a scientific theory T to be empirically grounded, it is not necessarily enough for there to exist *some* way of determining the values of the quantities it postulates. We are not entitled to say a theory T is satisfactorily grounded if the way of determining the values assumes that very theory T.

One natural way around the problem that might suggest itself would be to require that if a theory is to be empirically grounded it must be possible to determine the values of its properties without recourse to any theory at all. However, plausibly, for any theoretical, non-observational property, this requirement could not be met. *Some* theory must be used in determining the value of the property.

Following a suggestion due to Hermann Weyl, van Fraassen suggests that for a theory to be grounded it must be the case that a *number of different* theories yield the same values for the properties. Following Weyl, van Fraassen refers to this as the requirement of concordance. In summary, for van Fraassen, a theory is said to be "empirically grounded" if and only if it meets the conditions of determinability and concordance:

Determinability Any theoretically significant parameter must be such that there are conditions under which its value can be determined on the basis of measurement.

Concordance Comprising two aspects:

1. *Theory relativity*: this determination can, may and generally must be made on the basis of theoretically posited connections.
2. *Uniqueness*: the quantities must be "uniquely co-ordinated", there needs to be concordance in the values determined by different means.[34]

Van Fraassen argues that the work of Perrin ought not to be seen as making a case for realism about molecules, but rather as establishing that the theory of molecules is empirically grounded in the above sense.

In considering van Fraassen's claims, it is useful to distinguish between two questions: (i) Does Perrin's work establish the theory of molecules is empirically grounded? and (ii) Does Perrin's work *do nothing* more than this? We can accept, with van Fraassen, that Perrin's work does establish that the work of Perrin and others establishes concordance between a wide variety of ways of determining Avogadro's Number, and that this establishes that the theory of molecules is empirically grounded. But it need not follow that this is all that Perrin's work does. On the

[33] See van Fraaasen, *op cit*. p.8. Van Fraassen discusses the significance of an "Atwood Machine", and of a mechanism consisting of two masses joined by a spring, as devices for measuring mass.

[34] See van Frassen, *op cit*, p.11.

view defended here, it is not the fact of concordance, by itself, that is claimed to make realism about molecules rationally credible. Rather, it is concordance, together with a series of Eddington inferences and other probabilistically justifiable inferences that do this. We can agree that some aspects of Perrin's work establish that the theory of molecules is empirically grounded, but it has here been argued that the whole of his work is sufficient to give us good reason for saying molecules exist.

Stathis Psillos has replied to van Fraassen's arguments.[35] While agreeing with van Fraassen that Perrin's work does supply an empirical grounding for the theory of molecules, Psillos maintains that Perrin's work also supplies us with good reason for saying that molecules exist.

Psillos gives a Bayesian interpretation of Perrin's argument. More precisely, Psillos argues that some experimental work done prior to Perrin conferred upon the theory of molecules a reasonably high prior probability. Psillos argues that it follows from this that Perrin's results conferred upon the theory of molecules a high probability. He also argues that his approach does not commit the base rate fallacy.

It will be argued here that Psillos still does not get us all the way to realism. One shortcoming of the method of concordance, it has been argued, is that it fails to establish that Avogadro's Number of molecules, in the sense of tiny bits of matter, exist. The most it establishes, it has been argued, is that Avogadro's Number of *somethings* exist. It will be argued the same can be said of Psillos's approach.

As Psillos interprets Perrin, the latter makes the following two claims:

(i) If the atomic hypothesis is true, our experimental results ought to indicate the number of molecules in a mole of matter to be Avogadro's Number (6×10^{23}).
(ii) If the atomic hypothesis is not true, our experimental results could indicate the number of molecules in a mole of matter to be *any* number from zero to infinity. (On such an interpretation of the *negation* of the atomic hypothesis, it could turn out that the number of molecules in a mole *is* Avogadro's number, but the probability of this happening would be (as close as does not matter to) zero.)

Psillos argues that it follows from (i) and (ii), together the observed fact that many different methods agree that the number of molecules in a mole is Avogadro's Number, that the probability of the atomic hypothesis is very high.

However, it is surely far from clear that the truth of (ii) has been established. Let us assume "atomic theory" is construed as the claim that *tiny bits of matter*, *with mass*, *and moving around according to Newton's laws are responsible for the values we get when experimentally determining how many molecules there are in a mole*. If this is how we interpret "atomic theory", it would seem to be simply false that if atomic theory (in this sense) were wrong that our experiments could yield any value from zero to infinity for the number of molecules in a mole. As we have already noted, there could be 6×10^{23} somethings in a mole, without these somethings necessarily being tiny particles of matter. Perhaps, as Duhem suggested, *all* that we can

[35] See Stathis Psillos "The View from Within and the View from Above: Looking at van Fraassen's Perrin" in W. J. Gonzalez (ed), *Bas van Fraassen's Approach to Representation and Models in Science*, Synthese Library 368., (Springer, Dordrecht: 2014).

say about these "somethings" is that they are disposed to produce certain experimental results.[36] And if there were 6×10^{23} of these somethings – other than tiny bits of matter – it would seem our experiments would *still* yield a value of 6×10^{23} for Avogadro's Number. Consequently, if our experiments did yield a value of 6×10^{23}, we would not be entitled to conclude that this must have been due to molecules, considered as tiny particles of matter.

Against the foregoing it might be pointed out that the assumption that atoms or molecules are tiny masses moving around according to Newton's laws *plays an essential* role in many of the derivations used here. For example, it plats as essential role in the way Einstein arrives at his predictions about Brownian motion and in many of the other predictions tested by Perrin. If the assumption that atoms are tiny masses obeying Newton's laws *plays an essential role* in the derivation of these predictions, why do we also need Eddington inferences to reassure us atoms (in the sense of tiny masses obeying Newton's laws) really do exist?

Here it will be argued that we *do* need the Eddington inferences. One way of seeing this is as follows. Consider, for example, Einstein's demonstration that *if* there are tiny masses, obeying Newton's laws, affecting (say) the movement of suspended particles of gamboge, *then* the motions of the gamboge that would be produced by these tiny masses would have certain characteristics. And we do in fact find that the Brownian motion of the suspended particles of gamboge *has* these predicted characteristics. The logically important point to note here is that this is an "if…then…" statement: if there are tiny masses obeying Newton's laws then certain observable results will be obtained. And so, plainly, we would be guilty of the fallacy of affirming the consequent if we were to infer from such observations that these tiny masses do in fact exist. There might be other possible ways of explaining the same observations. According to (some versions of) the underdetermination thesis there *will be* other possible explanations. And Duhem has suggested another possible explanation: perhaps the entities responsible are not tiny masses obeying Newton's laws, but merely "somethings" that have a propensity to produce particular observational results.

What anti-Realists tend to deny is that the fact a particular derivation *requires us to postulate* some entities constitutes a sufficient reason for saying those entities exist. And so, to say that we are justified in saying atoms (in the sense of tiny masses obeying Newton's laws) exist since the derivations of Einstein and other require them, is to beg the question against the opponents of realism. Some additional reason is required for saying they exist, and this is what Eddington inferences purport to give.

Another reason might be given for saying that Eddington inferences are not required. Perrin established the concordance of many different methods. The theory that atoms and molecules really exist provides us with an explanation of this concordance. Further: the reality of atoms and molecules plausibly provides us with the

---

[36] See *Mixture and Chemical Combination and Related Essays* by Pierre Duhem, edited and translated with an Introduction by Paul Needham, *Boston Studies in the Philosophy of Science*, v 223, (Kluwer Academic Publishers, 2002), p.92.

best explanation of the concordance. This, it might be suggested constitutes a good reason for saying atoms and molecules are real.

But if this is given as the reason for accepting the theory of molecules as true, we once again find ourselves confronted with the question: "Is the fact that a theory is *the best* a sufficient reason to believe it is *true*?" And it has been argued we are not as yet in possession of a justification for saying it is true.

Psillos refers to some experimental findings that support the theory of molecules over other possible explanations of Brownian motion. In particular he refers to some work by Gouy which showed that Brownian motion was not due to convection currents.[37] But even if we accept that the work of Gouy ruled out convection currents, and the theory of molecules is the best theory we have still left standing, we surely have more argumentative work to do to show this best theory is probably true.

On the view advocated here, it is Eddington inferences that complete the job of getting us to the theory that it is *specifically molecules* that are responsible for the concordance. Eddington inferences establish that there probably exist particles of matter too small to see that obey Newton's Laws and hence also obey the Laplace Law of Atmospheres. They also establish that there probably exist tiny bits of matter too small to see that have the same random motion as Brownian motion. Perrin's work, together with Eddington inferences and the no-coincidental agreement inference, entitle us to conclude that it is the existence of *particles of matter too small to see* that are responsible for the fact that the different methods for determining Avogadro's Number are in agreement.

[37] Psillos, *op cit*. Gouy's findings are in L. Gouy "Le Mouvement Brownien et le Mouvement Moleculaires" in *Revue Generale des Sciences Pures et Appliquees*, v 6, pp.1–7.

Chapter 8
Eddington Inferences in Science – 2: The Size and Shape of the Universe

In the previous chapter we concerned ourselves with examples of "inward" Eddington inferences in science, that is, with inferences to claims about entities *smaller* than those we are capable of observing. But, of course, Eddington inferences can also take us "outwards", to claims about states of affairs larger than those we can observe. To refer once again to the example of the fish trap of Chap. 5, if we blindly set the holes of the trap to exactly four inches and get fish, we may infer that there are probably fish less than four inches in the sea, but we may with equal justification infer that there are also fish *longer* than four inches. In this chapter we examine some inferences that take us "outward" to things not observable by us because they are too big or distant. Some of these are straightforwardly Eddington inferences, while others are not. But all the inferences to be considered have the same underlying logical structure as those earlier considered: they begin by noting that the location of our observations is blindly chosen. They then note that, given that the location of our observations is blindly chosen, it would be a highly improbable fluke if some assertion S were not true. The conclusion is drawn that (probably) S is true.

8.1 Regions of Space and Time Outside the Observable Universe

Perhaps the most obvious example of an outward Eddington inference is the inference to the existence of regions of space-time that lie outside the observable universe. The observable universe consists of all those points of space-time such that there has been enough time since the big bang for light from them to have reached us by now. The observable universe in this sense is a (approximate) sphere, with us at its centre. Could we have good reason to believe there exist regions of space and time *outside* this sphere?

J. Wright, *An Epistemic Foundation for Scientific Realism*, Synthese Library 402, https://doi.org/10.1007/978-3-030-02218-1_8

A straightforward Eddington inference leads us to say there probably are such regions. Suppose it were the case that there was no space-time beyond the observable universe. Then a highly unlikely event would have occurred: the most distant point in space *visible to us* would have happened to have coincided with the most distant *existent* points. Since this is *a priori* highly unlikely, we may conclude that there probably are regions of space-time beyond the observable universe. And, of course, a similar argument assures us that those regions probably contain, for example, galaxies like those we can observe.

These conclusions seem to be generally accepted by cosmologists. But, as they stand, they do not say very much. They simply say that there (probably) exists *some* space, and *some* objects (probably galaxies), outside the observable universe. One of the aims of this chapter is to argue that there are purely probabilistic inferences leading us to say more specific things about these unobservable regions, and that, more generally, this is supportive of the thesis that we can have good reasons for scientific realist claims.

There is another point about the example just given that ought to be noted. It leads us to unobservable entities *without* any application of the No-coincidental agreement inference. This is because no claim is made that the entities to which we are led by an Eddington inference (unobservably remote galaxies) are identical with any entities postulated to *explain* some observable phenomena. And this is so generally with the examples considered in this chapter. Apart from a few possible cases, the status of which is controversial, our reasons for believing in things outside the observable universe do not come from any *causal explanation* such things may provide for observations we have made.[1] Rather, our reasons for believing in such things come solely from Eddington inferences and other probabilistically justifiable inferences.

8.2 Can We Make More Specific, Probabilistically Justified, Assertions About What Lies Beyond the Observable Universe?

In the previous section it was argued that we can have probabilistic reason for the claims that there exist regions of space time, and galaxies, further than we can see. But can we also have reason for claims that are more specific than the bare assertion

[1] Analysis of the motions of some objects at the periphery of the observable universe has been claimed to show those objects are moving in a way that cannot be accounted for solely in terms of the influence of other objects within the observable universe and the expansion of the universe. These (alleged) motions have been called "dark flow". Suggestions have been made that "dark flow", if it is exists, might be best explained by postulating a concentration of mass outside the observable universe, and even that it might be due to influence from another universe in the multiverse. However, the very existence of "dark flow" has been claimed to be dubious. See P. A. R. Ade et al. "Planck intermediate results XIII Constraints on peculiar velocities" http://arxiv.org/abs/1303.5090

that certain things *exist*? The main thesis to be defended in this chapter is that we can. More specifically, the thesis to be defended is that, subject to certain qualifications, Eddington and other probabilistic inferences can give us good reason for claims about the shape and size of the universe, including those parts of the universe unobservable to us.

8.3 Empirical Determination of the Curvature of Space

In this section we briefly review how empirical observations can furnish us with hypotheses about the shape of space.[2]

We begin by considering how an observer might do this when located on a two-dimensional surface; we then extend this to three dimensions. We will go through the inferential steps in arriving at a conclusion about the curvature of the surface in some detail to assure ourselves they need not rely on IBE.

Imagine an observer located on what might seem to be a flat plain. To the casual observer, it could *either* be truly flat, or curving slightly. Assume the observer can see that the plain seems to be randomly covered in stones. The observer could verify by observation that a number of areas (acres, say) on the surface are each covered in more or less the same number of stones. Perhaps each acre within the vicinity of the observer has more or less N stones. Such an observer could then use a variety of geometrical techniques to determine whether the region of the plain he can observe is flat or curved.

One technique appeals to the familiar geometrical fact that in flat space the inside angles of a triangle add to 180°. Our observer might lay out some strings of the surface of the plain, ensuring all points on the string were in contact with the surface, and the strings were straight in the sense that no shorter strings could be laid between the end points of each string. The observer could lay out three such strings to form a triangle. If the angles summed to 180°, the observer would have evidence the surface was flat, if they did not, there would be evidence it was not flat.

There are other techniques that could be used. The observer could draw out a circle from a central point, again using a stretched string. If the observer is on flat plain, the circle thus drawn will be larger than if the plain were curved as on the surface of a sphere.[3] Other possible curvatures of the plain would lead to larger

[2] The content of this section is elementary and introductory and many readers might wish to skip it.

[3] We can perhaps make this clearer by assuming that our observer is on a sphere, such as the Earth, and has traced out is a very large circle: in fact a "great circle" like the equator that divides the Earth in to two equal hemispheres. Then the distance from the central point to the circumference of the circle, as determined by pacing out that distance over the surface of the Earth, will be equal to one quarter of the circumference of the Earth. Say the distance from the centre of the Earth to the surface is 1. Then the circumference of the Earth is 2π. So, one quarter of this distance will be $\pi/2$, or about 1.57 units of length. That is, if we go from the central point (say, the north pole) to the equator by moving across the surface of the earth we cover 1.57 units of length.

circles.[4] Of course, the observer would need to determine the circumference of the circle, and there are a number of ways this might be done. One way this might be done would appeal to the fact that each acre within the vicinity of the observer had more or less N stones. And inductive inference leads to the conclusion that, probably, all acres have more or less N stones. Then, by counting the number of stones lying on, or within some distance of, the circumference of the circle, the observer could calculate the circumference. If the plain is flat there will probably be, say, M stones on the circumference. If the plain curved as the surface of a sphere, there will be fewer than M. And if the plain is "saddle shaped", there will be more.[5]

There is one more technique the observer could use. We have just noted that if the plain is flat there will be more stones on the circumference of the circle than there would be if the plain were curved like a sphere. This means that, on average, the angle subtended by adjacent pairs of stones on the circumference, as seen from the centre of the circle, will be greater if the plain is flat than they will be if the plain is curved like a sphere. If on a flat plain there would be M stones on the circumference, the average angle subtended would be, let us say, $n = 360/\mathrm{M}$. If our observer finds that the average angle is n, then there is evidence the plain is flat, if greater than n, evidence it is curved like a sphere, and less than n, that it is "saddle-shaped".

Now let us assume that our observer carries out the procedures described about, and finds that his observations suggest that he is on a curved surface. Perhaps he finds the number of rocks on the circumference to be less than M, or the average angle subtended greater than n. These observations *suggest* he is on a curved surface. But is the relation between the observations and the conclusion stronger than that? In particular, do the observations make it (more or less) *probable* that he is on a curved surface? Or is it instead the case that the hypothesis that he is on a curved surface provides merely an explanation (even if it is the best explanation) of the observations?

It might perhaps be suggested that the specific hypothesis that our observer is on a curved surface, like that of the Earth, is surely *only one possible* explanation of the obtained results. After all, there would seem to be many other shapes that could be attributed to the region around the circle, not to mention more remote regions beyond the circle, that could equally well account for the observations. This is of

---

Let us remind ourselves that – as noted two sentences back, the circumference of the Earth is 2π. The circle we arrive at (the equator) by traveling 1.57 units in a straight line from the pole, across the surface of the Earth, therefore has a circumference of 2π.

Now, let us assume that instead of being on the surface of the Earth, we are on a flat plain. We now imagine traveling out from our point of origin a distance of 1.57 units. But we are now traveling across a flat plain, not the curved surface of the Earth. Imagine a circle on this flat plain with a radius of 1.57. The circumference of this circle will, plainly, be $2\pi \times 1.57$. So, it will be a *bigger* circle than the one we form if we are on the curved surface of the Earth.

This is a simple geometrical fact. But this simple geometrical fact (or rather, the analogue of it in four dimensions) has a crucial role in enabling us to tell whether or not space is curved.

[4] A larger circle would be obtained if the plain were "saddle-shaped".

[5] This is discussed in more detail in Sect. 3, below.

course true, and so here we need to move carefully. Let us start by considering a rather more modest inference:

Premise (1): If our observer were on a perfectly flat plain, then the average angular distance between adjacent rocks on the circumference of the circle would be n.
Premise (2): The average angular distance between adjacent rocks on the circumference of the circle has been found to be greater than n.

Conclusion: The observer is not on a perfectly flat plan.

The conclusion is not that the observer *is* on a surface curved like the Earth, but rather the much weaker one that the surface is *not* perfectly flat. The argument is deductively valid: it is an instance of *modus tollens*. Premise (2), we may assume, has the status of (a mathematical consequence of) an observation and so is to be regarded as well founded. The only question to be considered is: do we have good reason to accept Premise (1)?

It might be protested that Premise (1) surely is subject to doubt. Perhaps the rocks aren't evenly spread across the plain – perhaps as we move away from the point at which the observer is standing, the rocks become more sparsely distributed. And if this were the case, the plain could still be flat even though the average angle was greater than n.

This, of course, is possible. But a *probabilistic* argument against this possibility can be given. The observer has established (by counting and measuring, let us say) that the rocks in his immediate vicinity – those located within some number of acres of his location – are distributed with a particular density. Then, on pain of his blindly chosen location being improbable, the most likely conclusion is that the rocks elsewhere are also distributed with more or less this density. The important point to note is that although we do, of course, have possible reason to doubt Premise (1), we can rebut that reason for doubt by appealing to purely probabilistic considerations.

Another reason might be given for doubting Premise (1). Suppose our observer has arrived at the conclusion that the average angular distance between all adjacent pairs of stones is greater than n, not by measuring each and every one of them, but only a sample of them. Then, it might be pointed out, Premise (1) – which asserts that this is true of all stones on the circumference – is thereby rendered subject to doubt.

But, again, a reply is available of the same form as that already used. Assume the observer has looked in one direction and found the average angle to be n. Might the average angle take some value in other directions? If so, an improbable event would have occurred: the blindly chosen *direction* for observation would have happened to have coincided with an "island" of stones with a lesser density in a sea of stones with a greater density. Moreover, probabilistic reasoning parallel to that already given assures us that the conclusion to be preferred is that, if measurements in other directions were to be taken, they too would probably say the density was the same as that observed.

So, probabilistic reasoning can be given in support of the truth of Premise (1).

There is one more possible reason for doubting Premise (1) that deserves slightly more extended discussion. It might be suggested that as we move away from the central point occupied by the observer: stones, rocks and even any measuring rods we might carry with us, systematically *expand* the further we go. A stone, for example, that might be 2 inches across at the centre of the observer's circle might steadily expand as carried away from this point. If everything – small grains of sand, rocks and boulders is systematically larger than it would be at the centre, then everything might, when surveyed by the observer by the methods described, present to the observer the appearance of a curved space and yet be flat.

There is one important difference between the possibility just sketched and the possibilities so far considered. Consider the possibility mentioned previously that the rocks are simply distributed more sparsely further out. Carrying a measuring rod out to the circumference of the circle and using it to determine the distance between the stones could, in principle, empirically test this. If there were a greater distance between the stones at the circumference than at the centre, this method would reveal this difference. But such a procedure would be ineffective in the scenario currently being considered. Any measuring rod carried out to the circumference would, in this scenario, itself expand, thereby rendering undetectable the increased distance between the stones.

However, although there would seem to be a strong sense in which this scenario would be empirically indistinguishable from the hypothesis of a curved surface, we can still have probabilistic reason to prefer the latter. It is a part of this imagined scenario that things steadily and systematically expand as they move away *from the location of the central observer*. We will say the location of the central observer is the *focus* of the expansions. But let us recall that the location of the observer is blindly chosen. The scenario under consideration requires us to say that the blindly chosen location of the observer happens to coincide with the focus of the expansions. But this is surely *a priori* highly unlikely. And so we have probabilistic reason to reject this scenario. But the hypothesis that the surface is curved does not require us to postulate any such highly improbable coincidental state of affairs.

So, it seems, given the data available to our observer, there is a probabilistic route to the conclusion that the surface on which he is standing is not flat. But, of course, to say it is merely not flat is much weaker than the conclusion it is curved like a sphere. And so we are now led to the question: Could our observer have probabilistic reason in favour of the stronger conclusion that it is curved like a sphere?

Let us begin by getting clearer on just what would be needed for an observer to be in possession of such evidence. We are assuming our observer has obtained evidence that at least to some extent, in some places, the surface on which he is standing is not flat. What needs to be done to show it is *likely* that he is on a sphere? Suppose, for simplicity, that the only observation our observer has made is that the angle subtended by two adjacent rocks X and Y on the circle is m, where $m > n$. Then, our observer will have evidence he is on a sphere if he has evidence that (1) the rate of curvature of the surface is the same in all *directions*, not just the direction of X and Y and (2) the rate of curvature is the same at all *distances* from the observer, not just the distance that X and Y lie from the observer. If the rate of curvature is the same at all distances and directions, it follows the observer must be on a sphere.

It will be argued that, actually, our observer already has enough empirical evidence to make the hypothesis that the entity on which he is standing is a sphere more likely than another specific hypothesis about its shape. The argument uses a mixture of inductive inferences and Eddington inferences. First, let us observe that it were not the case that the curvature is the same in all *directions*, then, improbably, the blindly chosen direction of observation would have happened to have coincided with an "island" of one specific rate of curvature of in a sea of other rates. This is unlikely, so by the argument of Chap. 2, the hypothesis to be preferred is that the curvature is the same in *all* directions.

Similarly, if the curvature were not the same at all distances, a less than maximally probable event would have occurred, and so the hypothesis to be preferred is that the surface on which the observer is standing has the same curvature at all distances from the observer.

It is important to note that this last conclusion does not quite get us to the conclusion that the observer is standing on a sphere. The conclusion of the last paragraph was that the surface has the same curvature at *all* distances. But, "all" has no "existential import". The conclusion tells us that if there are regions of the surface – the plain on which our observer is standing – beyond those that are visible to the observer, then they too will have the same curvature, but it does not say such regions exist.

To get the conclusion that other such regions exist, an Eddington inference is required. The inference can be represented as follows. Our observer occupies a blindly chosen location on the plain. From this blindly chosen location, the observer can see a region of the plain. Since these regions can be seen, we may say they exist. Do other regions, beyond that which can be seen, also exist? If they did not, a highly improbable fluke would have occurred: the blindly chosen location of the speaker, and therefore the blindly chosen limits of what the speaker can see, would have coincided with the boundaries of the plain. Since this is unlikely, we may conclude that there probably exist parts of the plain beyond those which can be seen. How far does the plain extend? From what has been argued earlier, restricted Eddington inferences could be given conferring some probability on the assertion the plain extends to any length. So, we have probabilistic reason to accept there exists indefinitely distant unobserved regions of the plain. But now, we have also established that, probably, all regions of the plain have the same curvature. And so it follows that there are indefinitely distant, unobservable regions of the plain that have the same curvature as the region on which the observer is standing. From these results it follows that the observer is, probably, on a sphere.

It might perhaps be wondered why an Eddington inference is required in our discussion of distance, but not direction. Recall that the function of Eddington inferences is to establish the probable truth of existence-claims. The observer cannot see beyond his horizon, and so an Eddington inference is needed to establish that there exist regions of the plain beyond the horizon. But, we may surely assume, our observer can look in all directions, and so he can establish by observation that the plain *exists* in all directions from his point of observation. If, hypothetically, our observer could not or had not looked in all directions, then an Eddington inference could be used to establish that, probably, there existed parts of the plain in directions the observer had not looked.

In summary, it has been argued that there is a purely probabilistic route from observations to the conclusion that our observer is on a sphere. These inferential moves confer an increased probability on a number of claims about parts of the sphere that cannot be seen by the observer from where he stands. These claims concern: the number of regions of a comparable area to the area he can see and the amount of curvature on each such area, the likely number of stones on each such area, and the likely area of the total sphere. Of course, the probability of these conclusions need not be very high, but it has been argued that they can be arrived at by purely probabilistic inferences, and do not rely on IBE.

8.4 Extending the Inferences from Two Dimensions to Three Dimensions, and to the Actual Universe

So far, we have only been concerned to investigate the ways in which a hypothetical observer on a two-dimensional surface might acquire purely probabilistic evidence that the surface was curved. But of course our aim is to determine whether or not it is possible to have purely probabilistic evidence for the existence and structure of regions of space beyond the observable universe. Can the inferences that we have applied to a two-dimensional surface also be applied to three-dimensional space?

In considering this question, it is useful to state clearly just what was argued in the previous section. It was argued:

(i) That there exist certain probabilistic inferences from initial data D, and these probabilistic inferences yield conclusions about the existence and shape of unseen parts of the surface on which the observer is standing.
(ii) The claims in D, that play the role of premises in the probabilistic inferences, are *observations*. (More specifically, they do not rely on IBE or any other questionable form of inference.)

We are considering whether the methods described in the previous section can be extended to inferences about what, if anything, might lie beyond the observable universe. Briefly, the position to be defended is that (i) is easily and naturally extended from two dimensions to three, but the status of (ii) is rather more complicated.

Let us begin by briefly sketching how the methods described in the previous section can be applied to three dimensions. In the previous section we saw that our observer would have evidence that the surface on which they were standing was curved if a circle on that surface had a smaller circumference than it would have had if the surface were flat. Applying this to three dimensions, an observer can have evidence that the space they are in is curved if a sphere has been found to have a smaller surface area than it would have if space were flat. It was also noted in the previous section that the observer would have evidence the surface was curved if distant object seemed larger – in the sense that its boundaries subtended a larger angle from the standpoint of the observer – than it would if the surface were flat. This same form of evidence is available when we move up to three dimensions: we have evidence that space is curved if distant objects appear larger, in the sense of subtending larger angles, than they would if space were flat.

In our example involving a two-dimensional surface, our evidence relied on the idea that the density of stones across the plain was probably more or less the same as in the immediate vicinity of the observer. The proposition that plays the analogous role in investigations in to the shape of the three-dimensional space of the actual universe is the Cosmological Principle:

There is (probably) nothing "special" about our position in the universe, and therefore (probably) nothing special about the position from which we are observing the universe. Consequently, at all points in space, the universe is (probably) more or less the same as it appears to us here and now.[6] ______________________________(CP).

---

[6] See Andrew Liddle *An Introduction to Modern Cosmology* (John Wiley and Sons, 2003), p.1–2. Liddle, *loc cit*, describes the cosmological principle as the "cornerstone" of modern cosmology.

Taken strictly and literally, (CP) would seem to be obviously false. The way things are here and now as I write are very different from the way they are in Antarctica or at the bottom of the ocean, and they are even more different in the centre of the Sun or in the frozen wastes of space. Still, in some form or another, the cosmological principle has had, and continues to have a role, particularly in using General Relativity to derive a model of the universe. The principle was used extensively by Einstein.

On the face of it, this presents us with a puzzle. Why should the principle have been used, and continue to be used, when it is *prima facie* obviously wrong? It will be argued that a natural explanation of this follows from the position developed here. There is an argument for the cosmological principle that closely parallels the argument for induction given in Chapter Two. This is of course a probabilistic argument and so its conclusion is defeasible. If it is defeasible, it may be falsified. And, in fact, the cosmological principle CP – as stated above – plausibly has been falsified. But although CP (on a strict interpretation) may have been falsified, the logic of the probabilistic argument in favour of some form of the cosmological principle retains its force. And so, it may yet be rational to accept some other restricted or modified version of the cosmological principle. Moreover, we find that this is exactly what scientists have done.

First let us see how an argument for the cosmological principle CP can be constructed that parallels the argument given in Chapter Two for induction.

> Suppose it were the case that volumes of space other than those observed contained significantly different properties from those that have been observed. If so, an improbable event would have occurred: the volume of space in which we, fortuitously, or by blind chance, find ourselves located, would have been an "island" with one set of properties in a sea of volumes with other sets of properties. Since this is *a priori* unlikely, we have reason to believe this is not the case, and therefore that other volumes probably exhibit more or less the same features as our own.

We will not here repeat the various arguments given in Chapter Two defending the thesis that "Other volumes are like our own" has a higher probability than any other specific claim about the nature of the other volumes.

It is clear that the type of inference defended in Chap. 2 can give probabilistic support to CP. If CP were not true, an improbable event would have occurred: the blindly chosen location of our point in space of observation would have happened to have coincided with a special or unusual position in the universe. The hypothesis that is more likely to be true than any other comparably specific hypothesis is that at all locations the universe exhibits more or less the same appearance as it does from our location in space.

One consequence of CP is:

However, as we have already noted, CP is false. The volume of space in this room as I write is very unlike an equivalent volume at the bottom of the ocean or the centre of the Sun. The inference to CP is not only defeasible but has in fact been defeated. However, scientists have not given up the cosmological principle altogether. Other, more restricted forms of the principle have been defended. While it seems to be clearly false that each room-sized volume of space is similar to each other room-sixed volume, perhaps larger volumes do resemble each other. After all, if we had a very large collection of marbles half of which were black and half white, it would not be hugely surprising if four marbles drawn at random were all white and another four all black. But, any million marbles drawn at random would be expected to closely resemble, in proportions of black and white, any other million drawn at random. Moreover, versions of the cosmological principle applying to larger volumes of space – more specifically, of the order of 250 million light years – seem to be as yet unfalsified and continue to be defended by astronomers. And this is in accord with the idea that there is a general, *a priori*, probabilistic but defeasible argument for some version of the cosmological principle. Even if certain versions of the principle prove to be false, this argument retains its force and so other versions of the principle are advanced and subjected to testing.

There is, moreover, a sense in which the version of the cosmological principle just mentioned – that is, that each sufficiently large volume of space is like every other sufficiently large volume – has been found to be false. It is *not* true for all volumes at all times through the history of the universe. According to the currently received view, there was once a time when all the matter in the universe was compressed down in to a very small volume: much smaller than a volume 250 light years across and in fact smaller than an atom. But still the principle is retained in a suitably qualified form: all sufficiently volumes of space at any one point in time contain (more or less) the same amount of matter as any other suitably large volume that exists at that point in time.

There are three more forms of the cosmological principle that are regarded as having stood up to testing. These are:

(i) The nomological form of the principle: The same laws of nature operate at all points of space and time throughout the universe.
(ii) The causal-historical form of the principle: All the entities and structures in the universe are the result of similar causal-historical processes. This is distinct from, and makes a more specific claim than, the nomological version of the principle. It asserts, for example, that even though galaxies in the distant past have very different properties from those around now, the ancient galaxies and the present ones are all evolving according the same causal-historical process of development. (This version of the principle might be regarded as the cosmological analogue of geological uniformitarianism.)
(iii) The isotropy form of the principle: the universe is held to exhibit broadly the same features no matter from which *direction* it is observed.

In summary, there is an *a priori* but defeasible probabilistic argument for the cosmological principle. Strong forms of the principle have been falsified. But the general, probabilistic argument for the principle remains and scientists have advanced more qualified versions of the principle, many of which remain unfalsified. All of this comports well with the view scientific inference defended in this book.

Matter is (probably) more or less homogeneously distributed throughout the universe. ________________________________(CPH)

CPH makes it possible for us to estimate the surface area of spheres in actual, three-dimensional space. If the number of objects, such as galaxies, lying on the surface of a sphere has fewer objects on it than would be the case if space were flat, then we have evidence the three-dimensional space we occupy is curved. Similarly, if the boundaries of distant objects were found to subtend larger angles than they would if space were flat, we would likewise have evidence our space is curved.

In the previous section it was noted that another assumption used in deriving the shape of the surface was that the features, such as curvature, exhibited by the surface were the same in all directions. The same principle can be used here:

The universe exhibits isotropy, that is, it exhibits more or less the same properties no matter from what direction it is observed.___(CPI).

Suppose we had observed that some distant objects subtended larger angles than we would expect to be the case if space were flat. Then, (CPI) would lead us to conclude that in all directions all such objects would subtend larger angles. This would constitute evidence that the surface of a large sphere surrounding us had a smaller area than would be the case if space were flat. Such a finding constitutes evidence that at least the space from our point of observation to the surface of the sphere is curved. An Eddington inference tells us that there probably exist volumes of space other than the one we can observe. Finally, CP tells us that the other volumes probably exhibit the same curvature as our own. These results would yield the conclusion we are probably living in a curved, closed three dimensional space with the properties of the three-dimensional surface of a four-dimensional sphere. Moreover, the degree of curvature of the space would enable us to estimate how far our universe extends beyond the parts we can see, and hence the size of the *unobservable* universe.

It is appropriate at this point to go into differences between flat and curved space in a little more detail, and to introduce some terminology. In a flat space, the inside angles of a triangle add to 180° and the circumference of a circle of radius r has the value of $2\pi r$ exactly. A flat surface is, naturally enough, said to have *no* curvature. The symbol denoting what is referred to as the curvature of space is conventionally denoted by the letter "k", and for a flat space, $k = 0$.[7] On a surface curved like a

[7] In cosmology, the value of k is given by the formula: $-2\ U/mc^2x^3$, where U is the total energy (kinetic energy and gravitational potential energy) of a mass m with respect to some other mass M, c is the speed of light, and x is the co-moving distance between m and the (centre of) the other mass M.

The co-moving distance between two objects can be explained as follows. It is now generally accepted that the universe is expanding, and this expansion is uniform, in the sense that the rate at which it is expanding is the same at all points within the universe. So, even if there are two objects (two stars, say) that are in all other respects at rest with respect to each other, the distance between them will be increasing merely in virtue of the expansion of the universe. Suppose, hypothetically, the two stars are initially one unit of distance apart. Then, in the course of some unit of time, the universe doubles in size. Then the two stars will now be two units of distance apart. Suppose after sphere, as we have noted, a circle (or if we move up a dimension, a sphere), with radius r will be smaller than one with the same radius on a flat surface. So, for such a space, the circumference of a circle will be less than $2\pi r$. And it is not too difficult to see that the inside angles of a triangle will come to over 180°. For such a space, $k > 0$. A universe like this is finite, and is said to be "closed". The other type of geometry a space can have is "hyperbolic". If a space is hyperbolic, then $k < 0$. In such a space, the circumference of a circle is greater than $2\pi r$ and the inside angles of a triangle come to less than 180°. A universe with this type of space is believed to be infinite, and is said to be "open".

In Sect. 8.1 we noted that Eddington inferences lend probabilistic weight to the conclusion that there are regions of space beyond the region we can observe. If $k = 0$, and space is flat, restricted Eddington inferences confer some probability on the conclusion that there are infinitely many such regions, and the space in those regions is also flat.[8] The cosmological principle CPH confers some probability on the claims that, if viewed, these regions would have a similar distribution of matter to the part of the universe we occupy, and CPI confers some probability on the conclusion that this is true of the distant regions in every direction from our own position. Corresponding conclusions can be drawn if $k < 0$.

Let us now consider the volumes of space that would (probably) lie outside the observable universe if $k \leq 0$. Such volumes would be "unobservable" in a *very* strong sense of the word. First, for any object outside the observable universe there has, by definition, not been enough time since the big bang for light from the object to reach us. Perhaps light from objects outside what is currently "the observable universe" might reach us at some point in the future. But if space then is flat or hyperbolic, then Eddington inferences lead us to say there will be no limit to how remote some objects are. By making the objects sufficiently remote, we can make any signal from such objects (that may arrive in the future) as weak as we please, and so undetectably weak. But, we still have *probabilistic* reason for saying they exist.

But there is also a stronger point that can be made. It is currently believed that our space is expanding. If so, and if $k \leq 0$ (or even if k is only slightly greater than

some longer span of time, the universe triples in size: the stars will now be three units of distance apart, and so on. So, the distance between the stars will, *as a matter of physical fact*, *have increased*. But, now let us imagine some grid or system of co-ordinates in space that itself expands at exactly the same rate as the universe expands. When the universe doubles in size, the spaces between the lines in the grid will also double in size, when the universe triples in size, so will the spaces between the lines of the grid. Clearly, relative to such an expanding grid, the distance between the stars will have remained the same. The *co-moving* distance between any two objects A and B can be thought of as the distance between A and B as measured by a grid or system of co-ordinates that expands (or contracts) exactly as the universe itself expands or contracts.

[8] This is something of an oversimplification. If the space of the observable universe is flat and Euclidean, then restricted Eddington inferences lead us to postulate indefinitely many other volumes of space beyond the observable universe that are also flat and Euclidean. However, it is possible for a space to be flat and non-Euclidean. For example, it is possible for a space to be flat and torus-shaped. Such a space would, plainly, be finite. However it seems to the present author that most cosmological discussions assume that the universe is not torus-shaped.

0), then there are some objects such that the light from them will never arrive at the Earth. Such objects are said to be beyond the *future visibility limit*. Any object more than 62 billion light years, or 19 gigaparsecs, from the Earth would be beyond this limit.[9] And if $k \leq 0$ – and the universe therefore probably infinite – there will probably be many such objects: in fact, almost *everything in the universe* will have this characteristic apart from the relatively infinitesimally small volume that is the observable universe.

8.5 Scientific Realism and the Unobservability of the Very Remote

Let us now consider what bearing the conclusions of the previous section may have to Scientific Realism. Consider the following assertion:

There exists a volume of space V, with curvature $k \leq 0$, containing matter of a similar density and kind that exists in the observable universe, but beyond the future visibility limit, and is therefore so remote from us that no signal from it of any sort can ever reach us.___________(UU)

It has been argued that we can have *probabilistic* reason for (UU).[10] But there is a clear sense in which V would be unobservable. Moreover, it seems clear that the sense in which V is "unobservable" is stronger than the sense in which molecules, atoms and even electrons are unobservable. There are numerous ways in which we can detect the effects of electrons. It is even seems to be the case that there is a sense in which we can obtain images of electrons.[11] But there is no possibility of observing V itself. No light or other signal from anything in V could ever reach us. And yet we have seen how we could have purely probabilistic reasons for saying that V and things in it *exist and have certain properties*. Our reasons for saying they exist need not rely on IBE. This would seem to be a particularly clear example of how we can

[9] See Matts Roos *Introduction to Cosmology* (John Wiley and Sons, 2015), p.218. A "gigaparsec" is 1 billion parsecs. A "parsec" is the distance from which one "astronomical unit" – the distance from the Earth to the sun, or about 93 million miles – would subtend an angle of 1/3600th of a degree.

[10] It is worth noting that the case, via an Eddington inference, for the existence of such a volume would seem to be pretty strong. The limit of the observable universe is about 14 gigaparsecs away. The future visibility limit is about 19 gigaparsecs away. So, a restricted Eddington inference does not need to be *very* restricted to get us to volumes of space beyond the future visibility limit: it only needs to postulate volumes of space about 1.35 times more distant than those we can see and therefore know exist.

A very high value for k, that is, a very strongly curving universe, could prevent there from being anything beyond the future visibility limit. However, later in the chapter it will be noted that the available empirical evidence mostly points to k having a much lower value than that which would be necessary to stop there being anything beyond the future visibility limit.

[11] See A. Stodolna et al. "Hydrogen Atoms under magnification: direct observation of the nodal structure of stark states" in *Physical Review Letters* **110**, 213001, May 2013.

have reasons for saying something exists without that thing being observable, and without relying on IBE.

It might perhaps be objected that the example just discussed is not properly comparable to the micro-entities (electrons and the like) that are usually doubted by the sceptics of scientific realism. The objects that are contained within unobservably distant volumes of space are, presumably, stars and galaxies and other objects of the same kinds as those found in the observable universe. Moreover, CP assures that entities like those in our vicinity are probably contained within those distant volumes. These distant entities, that is, differ from objects we have observed *only* in being more distant. But, it may be protested, the micro-entities that are rejected by instrumentalists and constructive empiricists are *not* like the more familiar objects that we can see. We do not regard electrons, for example, as being like observable objects such as basketballs or tennis balls. An electron is *not* "like a tennis ball, but much smaller": it is rather an entity of an entirely different sort altogether. So, the unobservable micro-entities do not differ from the entities we have observed merely in being less accessible: they also differ in being entities of a very different sort. And a sceptic about realism might say *this* is the reason why we do not have good reason to be realists about the likes of electrons. Even if it is true – the objection may continue – that we can have purely probabilistic reason for believing in the existence and general features of unobservably remote volumes of space, this does not in any way give aid or comfort to the scientific realist. The locus of disagreement between scientific realism and those who oppose it is rather different: it concerns the existence and properties of entities too small to see *and* that are *also quite unlike* the macroscopic entities with which we are ordinarily familiar.

It seems to me that this objection is not entirely accurate. Traditionally, there has been doubt about atoms, for example, that has arisen *solely* from their unobservability, and not from the fact that at the level of the extremely small, things start to display properties – quantum theoretic properties, for example – that seem to indicate that they may be very different from ordinary objects such as tennis balls, chairs, tables and the like. The sceptical doubts about atoms expressed by the contemporaries of Perrin, for example, seemed to have nothing to do with the "quantum weirdness" of atoms and molecules, but were rather entirely related to their lack of observability. If unobservability due to extreme smallness is seen as counting against realism, there is no evident reason why unobservability due to extreme remoteness ought not to be seen in the same way. And conversely, if it turns out that we can have probabilistic reason in support of claims about unobservably remote entities and states of affairs, then the case for realism would surely be strengthened.

Still, the feeling may persist that the example of stars that are unobservable because they are so remote is different from the example of atoms and molecules. Extremely remote stars are entities of the same sort as stars we can see in the following sense: If, hypothetically, our location in space were somehow moved to V we could see the stars in V. We could see that they look like, for example, the Sun. In this sense, the stars in V are *the same kind of thing* as stars that are observable, and *in this sense* are themselves "observable entities". But, it may be protested, atoms and molecules are not in this way the same kind of thing as observable entities.

However, one of the main themes of Chap. 5 was that this type of objection is mistaken. Let us return again to our fish trap. We could imagine our fish trap being moved around from one location to another, we could also imagine the size of the holes in it being made larger or smaller. Changing the location of the trap would change the regions of the sea from which the fish were caught, changing the size of the holes would change the size of the fish caught. Changing the trap in one way results in a change in the region from which the caught fish come, changing it in another way results in a change in the size of the fish. Fish in remote regions are uncatchable, given the blindly chosen *location* of the trap, smaller (or larger) fish are uncatchable given the blindly chosen size of the holes in the trap. And the same point could be made about human observers. Changing our blindly chosen location in space would result in a change in the location of the objects (such as stars) we could observe, changing our blindly chosen sensory apparatus may result in a change in the size of the things we could observe. Remote stars are unobservable given our blindly chosen actual location, objects too small to see are unobservable given our blindly chosen sensory apparatus. It is not entirely clear why the two sorts of case ought to be regarded differently.

It might perhaps be protested that changing our sensory apparatus would result in a change in "what it is" for something to count as "observable", but merely changing the location in which observation takes place has no such effect. This is perhaps plausible, but it is not clear what epistemological significance it might have. Given that the conditions of observation (whether they be location in space or the nature of our sensory apparatus) is blindly chosen, it is unlikely that the limits of what can be observed should coincide with the limits of what exists. So, we have probabilistic grounds saying things exist that we cannot observe, whether we cannot observe them because they are too far away or because they are too small. And this point remains even if changing the nature of our sensory apparatus would involve changing "what it is" for something to be observable.

There is another possible objection that perhaps needs to be briefly considered. It might be suggested that the example of unobservably remote stars and unobservably small objects are not relevantly similar because close stars and remote stars *are* objects of the same *natural kind*. But, it may be protested, macroscopic entities and the micro-entities of science are *not* objects of the same natural kind. For example, cats and electrons are distinct natural kinds of things, as are apples and hydrogen atoms. In the first type of case (close and distant stars) we are simply dealing with objects of the same natural kind that bear different relations to us, in the second type of case (cats and electrons, for example) we are dealing with objects of distinct, and very different, natural kinds.

One reply to this objection is that it is based on a premise that is, very plausibly, simply false. Macro-objects and unobservably small micro-entities *can* be of the same natural kind. It is not unusual for the claim to be made that a diamond is a very large molecule.[12] And the difference between a diamond and a (appropriately bonded) collection of carbon atoms that is too small to see is surely just a matter of size. They surely *are* objects of the same natural kind. The same claim has been made about silicon, graphite and the plastic polyethylene.[13] Things too small to see pretty clearly *can* be of the same natural kind as things large enough to see.

In summary, our discussion has supported the thesis that we can have good reasons for the types of entities and states of affairs the existence of which is claimed by scientific realists. It has been argued that we can have purely probabilistic reason for saying unobservably remote things exist. It has also been argued unobservably remote entities such as galaxies and stars ought to be seen as being in the same epistemic category as the unobservably small entities, such as atoms and electrons, which are perhaps more commonly the focus of debate between realists and their opponents.

8.6 Application to Actual Cosmology

We have not as yet considered the complexities of empirically determining the curvature and extent of the actual universe studied by cosmologists. It is useful here to remind ourselves that any inference to the features of the unobservable universe would have two components:

(i) A statement about certain features or empirically or observationally obtained aspects E of the *observable* universe.
(ii) An assertion that we are justified in making an inference from the claims about the *observable* universe E to the conclusion that some *unobservable* portion of the universe has features F.

It has been argued that the inferential move of (ii) can be given a probabilistic justification. *If*, for example, objects were sparser on the surface of a sphere some sufficiently large distance from Earth or seemed larger than they would if space were flat, we would have probabilistic reason to say space was curved. But do we in fact have good reason, that does not appeal to IBE, for making the claims about the required aspects of the *observable* universe? That is, do we, as a matter of empirical fact, have good, non-IBE dependent reason for (i)?

It is appropriate at this point to give a broad outline of some recent work. One technique for determining the curvature of space is based on observations of the Cosmic Microwave Background Radiation. This is frequently described as a “left-over” from the big bang. According to current theory, in its earliest stages the universe was a dense plasma through which photons were unable to travel any great distance. About 378,000 years after the big bang, the universe had cooled sufficiently

[12] See for example http://www.scienceline.ucsb.edu/getkey.php?key=530

[13] See for example Eng Wah Lim *Longman Effective Guide to O level Chemistry* (Pearson Education, 2007), p.56.

to enable protons and electrons to combine to form atoms of hydrogen. And this had the effect of enabling photons to now move freely through space. So, about 378,000 years after the big bang, the universe was suddenly filled with an immense "burst of light". This event is referred to as the photon decoupling. It is generally accepted that the photon decoupling is what is now picked up as Cosmic Microwave Background Radiation.[14]

There are a number of reasons why the photon decoupling is of use in measuring the curvature of space. The light from the photon decoupling is the oldest light (that has travelled any distance) that there is. In this sense, the photon decoupling is the oldest, and therefore most distant, visible event. It can be regarded as a very distant sphere, with us at the centre. This sphere is referred to as the "surface of last scattering".[15] But now, if our aim is to accurately measure the curvature of space, it seems desirable to use the largest sphere possible. So, for the purposes of measuring the curvature of space, the event of photon decoupling and the surface of last scattering would seem to be ideal.

There is another reason why the surface of last scattering seems like a good choice for measuring curvature. There is good reason to believe the intensity of the light throughout the surface of last scattering would *not* have been perfectly smooth or homogeneous. There must have been some inhomogeneities or "clumpings" in it: otherwise matter would never have coalesced in to stars and galaxies. These clumpings will show up as acoustic peaks in the Cosmic Microwave Background Radiation. According to the most widely accepted models, if space is flat, the average angular distance between these acoustic peaks in the CBR ought to be about one degree.[16]

This has provided astronomers with a way of measuring the curvature of space. If the average angular distance between the acoustic peaks is observed to be about one degree, we have evidence that space is flat and the curvature k of space is at least pretty close to zero. If the average angular distance is greater than one degree, we have evidence that $k > 0$, and that space has the features of a four-dimensional hypersphere. And if the angular distance is less than one degree, we have evidence $k < 0$ and that space is hyperbolic.

Observations of the anisotropies in the CBR have been obtained which have made it possible to calculate a value for k. The first experiment to do this used a

[14] This account comes largely from Liddle, *op cit*, pp.76–78.

[15] It is perhaps worth noting that this sphere is not a physical structure that would, for example, look like a sphere to any external observer. Rather, it is simply the set of all points that are a certain distance D from us such that any photons reaching us from that distance D will have been produced by the process of photon decoupling. Moreover, no photons from a greater distances – and therefore earlier times – could reach us since at those earlier times photons were unable to travel any great distance. So, the surface of last scattering is for us like a horizon when we look out at the universe. And just as different persons on different points on the Earth will have different horizons, so different observers in the universe (if there are any) will have different surfaces of last scattering. Also: just as the horizon on the earth for some observer on the surface of the Earth would not seem like a circular structure to an external observer, neither would our surface of last scattering.

[16] See, for example, http://www.star.le.ac.uk/nrt3/Cosmo/Cosmo11.pdf esp. p.12.

telescope suspended from a balloon high above Antarctica and is known as the "BOOMERanG" experiment.[17] The results of the experiment were found to most strongly confirm the hypothesis that $k = 0$, and that space is flat.[18]

Another feature of the early state of the universe that is useful in determining the curvature of space are structures producing "baryonic acoustic oscillations". The nature of the structures makes it possible to predict their size and hence the angle they would subtend when viewed from the Earth. Observations of the angles subtended by the structures, carried out by the BOOMERanG experiment, comport well with k having the value zero.[19]

Some very recent statistical analyses of the data, however, do point towards the possibility that space might have a very slight negative curvature. One recent study found k to have a value lying between +0.0011 and −0.0125, with a most likely value of −0.0057. The authors of the study concluded that, given the possibility of a small positive value for k, the minimum size of the universe, given the data, was 251 Hubble volumes. The maximum possible size, given the data, is infinitely large.[20]

8.7 Another Way of Measuring the Curvature of Space

The method of determining the curvature of space outlined in the previous section does so by ascertaining the geometrical features of space. It determines whether or not structures on the surface of last scattering seem from our position in space to as large as, bigger than, or smaller than they would seem to be if space were flat. There is, however, another, independent method of determining the curvature of space. This method relies on Einstein's General Theory of Relativity (GTR). According to GTR, the curvature of space depends on the amount of mass/energy contained within that space. A greater amount of mass/energy causes space to have a higher curvature, less mass/energy means a lower curvature.[21] This provides a way to predict the degree of curvature of actual space. First, according to GTR, there is a crucial amount of mass/energy that must be in the universe if space is to be flat. This crucial quantity is designated by the symbol ρ_c. If the amount of mass/energy in the universe is equal to ρ_c, then, according to GTR, space will be flat. If it is greater than

[17] See de Bernadis, P. et al. "A Flat Universe from High Resolution Maps of the Cosmic Microwave Background Radiation." *Nature*, **404**, (6781), (April, 2000), pp. 955–959. "BOOMERanG" is an acronym for "Balloon Observations Of Millimetric Extragalactic Radiation and Geophysics".

[18] See "Detecting the anisotropies in the CMB" in http://www.cambridge.org/au/download_file/192050/

[19] See "Detecting the anisotropies in the CMB" *loc cit*.

[20] See "Applications of Bayesian model averaging to the curvature and size of the universe" by M. Vardanyan, R. Trotta and J. Silk, *Monthly Notices of the Royal Astronomical Society*, **413**, L91-L95, (2011).

[21] See, for example, Liddle, *op cit*, p.50.

ρ_c then curvature will be positive and space will be closed. If it is less than ρ_c then space will have negative curvature and be open.

Current estimates show the actual quantity of mass energy in the universe to be very close to ρ_c and hence that space is probably (very close to being) flat. The total mass/energy in the universe is made up of the following four main components: ordinary matter (≈4.9%), "dark matter" (≈26.8%) and "dark energy" (≈68.3%). Ordinary matter is the matter that makes up bodies, such as stars and clouds of dust, that can be detected directly. "Dark matter" cannot be detected indirectly: the evidence for it is indirect. Some bodies can be observed to be moving through space in a way that suggests that they are under the influence of gravitational fields. The masses responsible for these gravitational fields cannot be seen, and so they are referred to as "dark matter". "Dark energy" is a form of energy postulated to explain the observed rate of expansion of the universe. Observations making it possible to calculate these quantities were recently obtained by the Planck space telescope.[22] When these forms of mass/energy are added up, the total is very close to that required by GTR for space to be flat.[23]

In summary, there are two ways of measuring the curvature of space within the observable universe. One way uses anisotropies in the CMB to determine the geometry of space. The other uses the General Theory of Relativity, together with estimates of the total mass/energy in the observable universe. Both ways give results indicating space is very close to being flat. Statistical analyses show that space is most likely open or flat, or, if closed, its curvature is very slight. Eddington inferences from these findings show it to be likely that that there exist volumes of space, and galaxies in them, that are in a very strong sense unobservable. It has also been argued this supports the realist claim that it is possible to have evidence for unobservable objects and states of affairs.

8.8 How Good Are the Foregoing Inferences?

We have seen that there are two different ways of measuring the curvature of space. But the two methods yield, to a fairly high degree of accuracy, the same result: both say the space of the observable universe is very close to flat.

Are we then entitled to say that the space of the observable universe *is* flat, and thereby have probabilistic grounds for the claims about unobservables referred to earlier? Unfortunately, there are a number of issues that need to be addressed before we are entitled to draw this conclusion.

First, it is clear that both methods for obtaining a value for the curvature of space are *highly* theory-dependent. The first method depends upon the predictions made by models of the early stages of the universe about the size of the anisotropies in the

[22] See Ade, P. A. R. et al. "Planck 2013 Results Papers" *Astronomy and Astrophysics*, **571**, A1, arXiv: 1303.5062.

[23] See Ade et al., *op cit*.

CMB. The second depends upon General Relativity, and on hypotheses concerning how certain observed motions within the universe, and the rate of expansion of the universe, are best explained. So, the methods used clearly rely upon theory, and upon IBE. However, this theory-dependence and reliance on IBE need not mean we aren't entitled to trust the methods. As we have already noted, we can have good reason to trust two methods *if they independently give us the same result.*

So, of course, we are now confronted with the question: are the two methods in fact independent? There is good reason for saying they are. Consider the first method. This uses observations of anisotropies or "unevenesses" in the surface of last scattering, a sphere at the limit of the parts of the universe visible from our position. The second method adds up the total/mass energy within the visible universe. This takes in to account the matter visible in stars, the mass required to explain motions of objects within the observable universe and rate of expansion of the universe. The two methods clearly rely on observations of different parts of the universe and phenomena. They are therefore independent.

However, even if it is granted that the two methods are independent, there are still possible objections that might be made to the claim we have good reason to trust them. In the previous chapter it was argued that there is a limit to just how much the agreement between independent methods ought to be taken as showing. It was argued that, although the concordance of different ways of determining Avogadro's Number made a good case for saying there are 6×10^{23} *somethings* in a mole, those "somethings" need not be molecules. Similarly, it might be asserted, the concordance of the different methods we have just outlined perhaps gives us reason to think that *something* has a value close to zero, but that something need not be the *curvature of space*.

In considering this considering this difficulty, it is worth noting that there are, of course, many possible ways of interpreting claims about space, and also many different views of what space is. Broadly speaking, though, we may say the following. Claims about space might be interpreted either instrumentally or, in a broad sense, "realistically". Interpreted instrumentally, claims about space are to be interpreted as the results of actual or possible measurements. Interpreted broadly "realistically", they are to be interpreted as something more than, or "over and above". the mere results of measurements: they are to be interpreted as something about the world. On a relationist conception of space, they are to be interpreted as claims about relations between things, on a fully realist or absolutist conception of space, they are to be interpreted as claims about space as an (immense) entity or object. It would take us too far beyond our concerns here to evaluate these different views of space. But, the following remarks can be made:

First, if we accept either of the broadly realist views as accounts of the space in the observable universe – that is, if we accept either the relationist conception of space or the fully realist, absolutist conception *as accounts of the space in the observable universe* – then the arguments given here furnish us with probabilistic reason for accepting assertions about volumes of space beyond the observable. If the fully realist absolutist conception is accepted about the space in the observable universe, then the arguments given here give reason for the existence of such fully realist but strongly unobservable volumes of space beyond the observable universe. And if the relationist conception of space is accepted about the observable universe, then the arguments give reason for strongly unobservable systems of relations beyond the observable universe.

Even if some form of operationalist or instrumentalist view of space is adopted, the arguments given here still lead us to *some* form of realism. We are, for example, led to make claims about what the outcomes would be of possible strongly unobservable operations or measurements. It must be conceded this is a long way short of what is usually thought of as realism, but is still in at least one sense realist: it says we can have good reason for the features of certain strongly unobservable states of affairs.

Let us now return to the difficulty with which we are concerned. The difficulty is that the concordance of different ways of measuring the curvature of space leave it undetermined just what it is that has curvature. However, it has just been argued that, no matter precisely what it is that is claimed to have curvature, the inferences described above support some sort of realism about the existence of strongly unobservable objects or events: whether it be unobservable regions of space considered as an entity, or as a system of relations, or about the results of strongly unobservable measuring operations. The inference to *some* sort of realism is retained even if concordance leaves it underdetermined just what it is that has curvature.

It is also worth noting that, of course, the inferences to the existence of distant, strongly unobservable galaxies and other physical objects remains intact. The inferences defended here support realism about at least some strongly unobservable things.

8.9 Further Uses of Eddington Inferences

The main examples of Eddington inferences used here concern the inference to atoms and molecules – in the sense of pieces of matter too small to see that are responsible for phenomena associated with heat – and the inference to unobservably remote regions of space, their curvature, and the entities (stars, galaxies) in them. On the face of it, Eddington inferences would only seem to be able to lead us to say some class of unobservable entities exist if those entities possess properties that are possessed by some *observable* things. But, of course, many of the properties attributed to unobservable entities are quite unlike the properties we can establish observable entities to possess. For example, electrons are held to have the property of spin, but observable objects like tables and chairs do not have "spin" in this sense. So, Eddington inferences would not seem likely to be able to help us show unobservables with spin exist.

Nonetheless, it will be argued here that Eddington inferences may still have some role in helping to establish the existence of unobservable entities with properties not apparently directly rooted in experience.

We have in fact already considered one such example earlier in the book. Consider visible light. Visible light with the lowest frequency is red, that with the highest frequency is violet. Is it reasonable to believe that there are forms of light with frequencies higher or lower than that which is visible? An Eddington inference shows that there probably are. Suppose that there was no light beyond the visible. Then, a highly improbable fluke would have occurred: the only frequencies of light that actually exist happen to coincide with the frequencies detectable by the human eye. But since this would be an *a priori* improbable fluke, we may conclude that probably there are forms of light with higher and lower frequencies. That is, infra-red and ultra-violet light probably exist. But we may go further than this. It follows from the position defended in Chap. 5 that we would have some probabilistic reason to believe that there exist frequencies of light indefinitely higher than the visible. So, we would have some probabilistic reason to believe in, for example, the existence of gamma rays.

Of course, no claim is made here that we can derive the claim "Gamma Rays exist" by an Eddington inference from premises that can be established by observation. To carry out such an inference, we would at least need some premise such as "Visible light is composed of *waves*". And it has not been argued here that Eddington inferences can establish this. Rather, the point of the example is merely to illustrate another way in which Eddington inferences could, under certain circumstances, carry us to claims about unobservable states of affairs.

Another other example of an Eddington inference, applied to wholly theoretical properties, comes from the branch of modern physics known as quark theory. The physicist Murray Gell-Man developed a theory that organised mesons and spin ½ baryons in to octets. It was a natural consequence of Gell-Man's system of organisation that the spin 3/2 baryons ought to form a decuplet.[24] However, at the time that Gell-Man postulated his theory, only nine spin 3/2 baryons were known. Gell-Man postulated a tenth such baryon, which he called the Ω^-. Subsequent experimental work confirmed the existence of the particle.[25]

Our concern here is with the broad features of the reasoning that led Gell-Man to postulate the existence of the new particle. It can be represented as an Eddington inference. Gell-Man had found all known mesons, spin ½ baryons and spin 3/2 baryons to conform to a particular mathematical pattern or structure. But he also noted that there was a "gap" in the actually known instances of the structure. There could be as many as ten 3/2 baryons exhibiting the general structure Gell-Man had discerned. But only nine such particles had been observed. Is it the case that those nine *so far observed* particles were the only ones that *actually existed*? If so, a somewhat improbable fluke would have occurred: the parts of the pattern we had observed would have happened to have coincided with the only parts of the pattern that were, in the actual world, instantiated. But since this seems unlikely, we can conclude this

[24] See M. Gell-Man "Symmetries of Baryons and Mesons" in *Physical Review Letters*, **125**, 1067, (1962).

[25] See V. E. Barnes et al., "Observation of a Hyperon With Strangeness Minus Three"., *Physical Review Letters*, **12**, (8): 204. (1964).

is probably not the case, and therefore that the so far unobserved part of the pattern – the Ω^- – probably also exists. An Eddington inference led Gell-Man to postulate the Ω^-.

Of course, it is not here asserted that we can be led by Eddington inferences *purely from observational data* to the Ω^- particle. The inference requires us to have already established the existence of other 3/2 baryons. And the question whether their existence can be established by Eddington inferences, or in any other way, has not here been considered. The point of the example is simply to illustrate how, once the existence of a class of entities has somehow been established, Eddington inferences can furnish us with good reason for saying that so far unobserved entities, bearing certain relations to the established entities, probably do exist.

8.10 Quantum Theory

In the previous chapter it was argued Eddington inferences help to establish the existence of atoms and molecules. But of course physical theory has since gone much further in to the nature of matter than that. And so the question arises: can the approach advocated here lead us any "deeper" than atoms and molecules? In particular, can it be used to justify some sort of realism about the realm described by quantum theory? In this book, no claim is made that it can. Ought this to be seen as a shortcoming of the view advocated here? It will be argued that it need not be seen as being so.

First, no claim is made that the route to realism comprising Eddington inferences and the other (probabilistically justifiable) inferences described here is the *only* possible way that realist claims of any sort can be justified. It has here only been argued that the approach furnishes us with *one* way to realist theses, and that *to date* no other purely probabilistic route to realism has been developed. It is compatible with the view adopted here that such other routes may exist. This view is compatible with realism about quantum theory.

There is another point that deserved to be made. It is surely fair to say that it still highly controversial just how quantum theory ought to be interpreted. Is it to be given a realistic interpretation at all, and, if so, what is the nature of that interpretation? This is still surely a matter of controversy. What, if anything, ought we to be realistic *about* when it comes to quantum theory? One reasonable answer is: We do not know.

This has relevance for any discussion of the epistemological basis of scientific realism. Our aim has been to supply a view of that which constitutes good reasons for scientific realist claims. Now, plausibly, *if* we do in fact have good grounds for making some particular realist claim P, then it is *ceteris paribus* desirable that our view be able to explain in virtue of what we do have such good grounds. But if we do not have good grounds for P, or if it is unclear that we do so, then it would appear to be no shortcoming of our view if it did not supply us with an account of those (possibly existing, possibly not existing) reasons. And this would plausibly seem to the situation we find ourselves in with quantum theory. Plausibly, we do not know what, if anything, we ought to be realists about when it comes to quantum theory. And so it ought to be seen as no shortcoming of the view offered here that it does not furnish us with reasons for realism in that area.

8.11 Is the Method of Eddington Inferences Too Limited?: Eddington Inferences and IBE Again

It has been argued that the view advocated here justifies us in making realist claims in a variety of areas, but the feeling might persist it justifies realism in rather fewer areas than we need. This can be brought out by considering some examples from physical chemistry. The properties of DNA and RNA, for example, go beyond those of tiny masses obeying Newton's laws. And so, the question arises: what stance ought to be adopted to those properties of DNA and RNA? More specifically, if their existence cannot be established by Eddington inferences, does it follow that we ought to refrain from asserting they exist? And if we *do* refrain from asserting their existence, will we still be able to explain the behaviour of DNA and RNA? Perhaps we would be left *unable* to explain them.

This might be thought to give us reason to prefer IBE to the view advocated here. After all, if we rely on IBE then obviously we *will* be able to explain the behaviour of DNA, or RNA, or of any subject matter at all, provided we have a (sufficiently good) theory of that subject matter. Relying on Eddington inferences might therefore be thought to restrict our explanatory abilities in a way that IBE does not. And if that is so, it might be wondered, is it really a good idea to privilege Eddington inferences over IBE?

In addressing this matter, let us first consider the suggestion that the view adopted here would leave us without a way of explaining, for example, the behaviour of DNA. It will be argued this is not quite right. There is obviously a sense in which the existing explanations will still be *available*, it is just that on the position adopted here it may be more appropriate to view them merely as *possible explanations*, rather than as explanations that have been shown to be more likely than their rivals.

Still, it may be felt that there is an advantage to using IBE. One the view advocated here, scientific theories that cannot be appropriately supported by Eddington inferences are reduced to the status of mere possible explanations. And it might be felt that IBE remains more attractive if it can confer a higher status than mere "possible explanations" on a wide range of our theories.

However, I think closer examination reveals this supposed advantage of IBE to be not genuine. The fundamental point is that it is not enough for some philosophical position to simply *have as one of its consequences* that we are entitled to say, for example, that a theory about the workings of DNA is true. It needs to be shown we have the *epistemic right* to say the theory is true. More specifically, proponents of IBE need to be able to show we have the epistemic right to assert that our theories about unobservables are true. And here it has been argued that proponents of IBE have not, as yet, succeeded in showing this. The supposed advantage of IBE is not genuine.

8.12 Concluding Remarks

The aim of this book has been to provide probabilistically justifiable, albeit defeasible, foundations for inferences to unobservable entities. In earlier chapters it was argued that we do not possess a satisfactory justification of IBE, where the best explanation is taken to be the explanation with greatest simplicity, or some kindred notion. It was also argued that there is some, limited, force to the Pessimistic Meta-Induction: A significant number of the unobservable entities postulated by past best theories have subsequently turned out to not exist. It was concluded that a route to unobservable entities that does not rely on IBE is needed.

The central concept that has been used is that of the "Eddington inference". It has been argued that inferences of this sort can be given an *a priori*, defeasible justification like that which can be given for induction. Therefore, some inferences to theoretical entities can be given a purely probabilistic justification that does not rely on IBE.

It has also been argued the approach applies to cases from the history of science, and to recent and contemporary science. More specifically, it was argued the approach can apply to inferences to atoms and molecules, and to unobservably remote regions of space.

On the view advocated here the epistemic status of existence claims about at least some unobservable entities is broadly of the same category as that of ordinary inductive predictions like "The Sun will rise tomorrow." The justifications for the two classes of statements are broadly similar. But since the belief the sun will rise tomorrow is surely a paradigm of rational acceptance, and some claims about unobservables enjoy a similar epistemic status, it follows that inferences to unobservable entities can be rationally justifiable. The view advocated here therefore meets our Requirement of Epistemic Adequacy: it establishes that at least some scientific realist claims are worthy of rational belief.

Bibliography

Abdelkader, M. (1983). A geocosmos: Mapping outer space in to a Hollow Earth. *Speculations in Science and Technology, 6*, 81–89.

Achinstein, P. (2001). *The book of evidence*. New York: Oxford University Press.

Ade, A. R. (2013). Planck 2013 research papers. *Astronomy and Astrophysics*, 571, A1, arxiv1303.5062.

Ade, A. R. (2014). *Planck intermediate results – XIII. Constraints on peculiar velocities*. http://www.arxiv.org/abs/1303.5090

Akaike, H. (1971). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov and F. Csaki (1971) (pp. 67–281).

Alai, M. (2014a). Deployment versus discriminatory realism. In *New thinking about scientific realism*. http://www.philsci-archive.pitt.edu/10551/

Alai, M. (2014b). Review: Explaining science's success. *Metascience, 23*, 125–130.

Barnes, V. E. (1962). Observations of a hyperon with strangeness minus three. *Physical Review Letters, 12*(8), 204.

Barrow, J. D. (1990) The mysterious lore of large numbers. In Bertolloti (1990).

Beardsley, M. (1958). *Aesthetics: Problems in the philosophy of criticism*. New York: Harcourt, Brace and World.

Bergstrom, L. (1984). Underdetermination and realism. *Erkenntnis, 21*, 349–365.

Bergstrom, L. (1993). Quine, underdetermination and skepticism. *Journal of Philosophy, 90*, 331–358.

Bertolotti, B., et al. (Eds.). (1990). *Modern cosmology in retrospect*. Cambridge: Cambridge University Press.

Blackmore, J. (1995). *Ludwig Boltzmann: His later life and philosophy: Book two*. Dordrecht/London: Kluwer Academic Publishers.

Bohr, N. (1913). On the constitution of atoms and molecules. *Philosophical Magazine, 26*(151), 1–24.

BonJour, L. (1980). Externalist theories of empirical knowledge. *Midwest Studies in Philosophy, 5*, 53–57.

BonJour, L. (1998). *In defence of pure reason*. New York: Cambridge University Press.

Born, M., & Wolf, E. (1989). *Principles of optics: Electromagnetic theory of propagation, interference and diffraction of light* (p. 986). Cambridge: Cambridge University Press.

Boyd, R. (1984). On the current status of the issue of scientific realism. *Erkenntnis, 19*, 45–97.

Carnot, S. (1824). *Reflections on the motive power of heat*. Paris: Bachelier.

Chakravartty, A. (2007). *A metaphysics for scientific realism: Knowing the unobservable*. Cambridge: Cambridge University Press.

Chalmers, A. (1976). *What is this thing called science?* Indianapolis: Queensland University Press.

J. Wright, *An Epistemic Foundation for Scientific Realism*, Synthese Library 402, https://doi.org/10.1007/978-3-030-02218-1

Chandler, J. (2013). Contrastive confirmation: Some competing accounts. *Synthese, 190*, 129–138.
Chang, H. (2003). Preservative realism and its discontents: Revisiting caloric. *Philosophy of Science, 74*, 902–912.
Clarke, S., & Lyons, T. (2002). *Recent themes in the philosophy of science*. Dordrecht: Kluwer Academic Publishers.
Clendinnen, F. J. (1982). Rational expectation and simplicity. *What? Where? When? Why?: Australasian Studies in the History and Philosophy of Science, 1*, 1–25.
Collins, H. (1985). *Changing order: Replication and induction in scientific practice*. Chicago: University of Chicago Press.
de Bernadis, P., et al. (2000). A flat Universe from high resolution maps of the cosmic microwave background radiation. *Nature, 404*, 6781.
Derske, W. (1992). *On simplicity and elegance*. Delft: Eburon Press.
Doppelt, G. (2005). Empirical success or explanatory success: What does current scientific realism need to explain? *Philosophy of Science, 72*, 1076–1087.
Doppelt, G. (2007). Reconstructing scientific realism to rebut the pessimistic meta-induction. *Philosophy of Science, 74*, 96–118.
Duhem, P. (2002). *Mixture and chemical combination and related essays.* Edited and translated with an Introduction by Paul Needham in *Boston Studies in the Philosophy of Science* vol. 223 Kluwer Academic Publishers.
Eddington, A. E. (1938). *The philosophy of the physical sciences: The Tarner lectures*. Cambridge: Cambridge University Press.
Egg, M. (2013). Review: Explaining science's success. *Dialectica, 67*, 367–372.
Einstein, A. (1956). On the movement of small particles suspended in a stationary liquid demanded by the molecular kinetic theory of heat. In R. Furth (1956) (pp. 1–18).
French, S., & Kamminga, H. (1993). *Correspondence, invariance and heuristics*. Dordrecht: Kluwer Publishing Company.
Friedman, M. (1974). Explanation and scientific understanding. *The Journal of Philosophy, 71*, 5–19.
Furth, R. (1956). *Investigations on the theory of the brownian movement*. New York: Dover Publications.
Garber, E., Brush, S., & Everett, C. (1986). *Maxwell on molecules and gases*. Cambridge, MA: MIT Press.
Gell-Man, M. (1962). Symmetries of baryons and mesons. *Physical Review Letters, 125*, 1067.
Gettier, E. (1963). Is justified true belief knowledge? *Analysis, 23*(6), 121–123.
Ghins, M. (2002). Putnam's no-miracle argument: A critique. In Lyons, & Clarke (2002) (pp. 63–90).
Goldman, A. (1986). *Epistemology and cognition*. Cambridge, MA: Harvard University Press.
Gonzales, W. J. (Ed.). (2014). *Bas van Fraassen's approach to representation and models in science*, Synthese Library. Dordrecht: Springer
Goodman, N. (1954). *Fact, fiction and forecast*. Cambridge: Harvard University Press.
Gouy, L. (1888). Le Mouvement Brownien et le Mouvement Moleculaires. *Revue Generale des Sciences Pures et Appliquees, 6*, 1–7.
Hahn, L., & Schilpp, P. (1986). *The Philosophy of W. V. Quine*. Chicago: Open Court Publishing.
Hajek, A. (2012) Interpretations of probability. In *The stanford encyclopaedia of philosophy* edited by Edward N. Zalta.
Harman, G. (1965). The inference to the best explanation. *The Philosophical Review, 74*, 88–95.
Hecht, E. (2014). *Optics*. Essex: Pearson Education Limited.
Hempel, C. (1945). Studies in the logic of confirmation. *Mind*, 54, 1–26, 97–121.
Howson, C. (1973). Must the logical probability of a theory be zero? *British Journal for the Philosophy of Science, 24*(2), 153–163.
Howson, C. (1976). *Method and appraisal in the physical sciences*. Cambridge: Cambridge University Press.
Hume, D. (1748) *An enquiry concerning human understanding*. Chicago, Encyclopædia Britannica

Hutchison, K. (2002). Miracle or mystery: Hypotheses and predictions in Rankine's thermodynamics. In S. Clarke, & T. Lyons (2002) (pp. 91–120).
Jeffreys, H. (1961). *Theory of probability*. Oxford: Clarendon Press.
Jones, R. (1991). Realism about what? *Philosophy of Science, 58*(2), 185–202.
Joule, J. P. (1854) On the existence of an equivalent relation between heat and the ordinary forms of mechanical power. *Philosophical Magazine*, 3, 27, 205–207.
Kennefick, D. (2007). Not only because of theory: Dyson, Eddington and the competing myths of the 1919 eclipse expedition. Cornell University Library. Available at http://arxiv.org/abs/0709.0685
Kitcher, P. (1993). *The advancement of science*. Oxford: Oxford University Press.
Keynes, J. M. (1921). *A treatise of probability*. London: MacMillan and Co.
Kostenbauer, S. (Ed.). (2008). *Pre-Proceedings of the 26th international Wittgenstein symposium*. Kirchburg am Wechsel: Wittgenstein Society.
Kukla, A. (1993). Empirical equivalence and underdetermination. *Analysis, 53*, 1–7.
Kyburg, H. (1970). More on maximal specificity. *Philosophy of Science, 37*, 295–300.
Ladyman, J. (2011). Structural realism versus standard scientific realism: The case of phlogiston and de-phlogisticated air. *Synthese, 180*, 87–101.
Ladyman, J. (2014). Structural realism. In E. N. Zalta (Ed.), *The stanford encyclopaedia of philosophy.* http://www.plato.stanford.edu/archives/spr2014/entries/structural-realism
Laudan, L. (1981). A confutation of scientific realism. *Philosophy of Science, 48*, 19–48.
Laudan, L., & Leplin, J. (1991). Empirical Equivalence and Underdetermination. *Journal of Philosophy, 88*, 449–472.
Lehrer, K., & Cohen, S. (1983). Justification, truth and coherence. *Synthese, 55*, 191–207.
Leplin, J. (1984). *Scientific Realism*. New York: University of California Press.
Leplin, J. (1997). *A novel argument for scientific realism*. New York: Oxford University Press.
Leplin, J. (2014). Review: Explaining science's success. *Analysis, 74*, 184–185.
Levin, M. (1984). What kind of explanation is truth? In Leplin (1984).
Lewis, D. (1973). *Counterfactuals*. Oxford: Basil Blackwell.
Liddle, A. (2003). *An introduction to modern cosmology*. Chichester: Wiley.
Lim, E. W. (2007). *Longman effective guide to o-level chemistry*. London: Pearson Education.
Lyons, T. (2002) Scientific realism and the pessimistic meta-*modus tollens.* In S. Clarke, & T. Lyons (2002) (pp. 63–90).
Maxwell, J. C. (1986). Molecules. In Garber, Brush and Everett (1986).
McDowell, J. (1978). Physicalism and primitive denotation: Field on Tarski. *Erkenntnis, 13*, 131–152.
Meixner, J. & Fuller, G. (2008). BonJour's *A Priori* Justification of Induction. In S. Kostenbauer (Ed.) (2008) (pp. 227–229).
Moore, G. E. (1959) Proof of an external world. In his *Philosophical* papers. (Routledge) esp. pp. 126–148.
Musgrave, A. (1976). *Why did oxygen supplant phlogiston?* In C. Howson (1976) (pp. 181–210).
Musgrave, A. (1988). *The ultimate argument for scientific realism*. In R. Nola (Ed.) (1988).
Musgrave, A. (1992). Realism about what? *Philosophy of Science, 59*, 691–697.
Newburgh, R., Peidle, J., & Rueckner, W. (2006). Einstein, Perrin and Avogadro's number – 1905 revisited. *American Journal of Physics, 74*, 478.
Newton, I. (1729). *Mathematical principles of natural philosophy* (A. Motte, Trans.). London: Benjamin Motte
Niiniluoto, I. (1987). *Truthlikeness*. Dordrecht: Reidel.
Nola, R. (1988). *Relativism and realism in science*. Dordrecht: Kluwer Academic Publishers.
Oddie, G. (1986). *Likeness to truth*. Dordrecht: Reidel.
Perrin, J. (1916). *Atoms*. London: Constable and Company.
Petrov, B. N. & Csaki, F. (1971). 2nd international symposium on information theory (Akademiai Kiado).
Poincare, H. (1905). *Science and hypothesis*. New York: The Walter Scott Publishing Co..

Popper, K. (1963). *Conjectures and refutations*. London: Routledge and Kegan Paul.
Popper, K. (2002). *The logic of scientific discovery*. London: Routledge.
Psillos, S. (1994). A philosophical study of the transition from the caloric theory to thermodynamics: Resisting the pessimistic meta-induction. *Studies in the History and Philosophy of Science, 25*, 159–190.
Psillos, S. (1999). *Scientific realism: How science tracks the truth*. London: Routledge.
Psillos, S. (2014). The view from within and the view from above: Looking at van Fraassen's Perrin. In W. J. Gonzales (Ed.) (2014).
Putnam, H. (1975). *Matter, mathematics and method*. Cambridge: Cambridge University Press.
Putnam, H. (1978). *Meaning and the moral sciences*. Abingdon: Routledge/Kegan Paul.
Putnam, H. (1981). *Reason, truth and history*. Cambridge: Cambridge University Press.
Quine, W. V. (1960). *Word and object*. Cambridge, MA: MIT Press.
Quine, W. V. (1970). On the Reasons for the Indeterminacy of Translation. *Journal of Philosophy, 67*, 178–183.
Quine, W. V. (1975). On empirically equivalent systems of the world. *Erkenntnis, 9*, 313–328.
Quine, W. V., & Ullian, J. S. (1978). *The web of belief*. McGraw Hill.
Rankine, W. J. M. (1853). On the mechanical action of heat, especially in gases and vapours. *Transactions of the Royal Society of Edinburgh, 20*, 147–190.
Ritchie, A. (1984). *Reflections on the philosophy of Arthur Eddington*. Cambridge: Cambridge University Press.
Roos, M. (2015). *Introduction to cosmology*. Chichester: Wiley.
Rutherford, E.. (1911). The Scattering of α and β particles by matter and the structure of the atom. In the *Philosophical Magazine*. Series 6, Vol. 21, May edition.
Salmon, W. (1984). *Scientific explanation and the causal structure of the world*. Princeton: Princeton University Press.
Salmon, W. (2005). *Reality and rationality* (edited by P. Dowe and M. Salmon). Oxford University Press.
Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 2*, 451–464.
Skyrms, B. (2000). *Choice and chance*. Cengage Learning: Wadsworth.
Smart, J. J. C. (1963) *Philosophy and* scientific r*ealism.* Routledge.
Smart, J. J. C. (1986). Quine on space-time. In Hahn, & Schilpp (1986).
Stanford, P. K. (2006). *Exceeding our grasp: Science, history and the problem of unconceived alternatives*. Oxford: Oxford University Press.
Stodolna, A., et al. (2013). Hydrogen atoms under magnification: Direct observation of the nodal structure of Stark states. *Physical Review Letters, 110*, 213001.
Stove, D. (1982). *Popper and after: Four modern irrationalists*. Oxford: Pergamon Press.
Stove, D. (1986). *The rationality of induction*. Oxford: Clarendon Press.
Thomson, J. J. (1904) On the structure of the atom. *Philosophical Magazine*, Series 6, **7**(39), 237.
Van Fraassen, B. (1989). *Laws and symmetry*. Oxford: Clarendon Press.
Van Fraassen, B. (2009). The perils of Perrin: In the hands of philosophers. *Philosophical Studies, 143*, 5–24.
Vardanyan, M., Trotta, R., & Silk, J. (2011). Applications of Bayesian model averaging to the curvature and size of the Universe. *Monthly Notices of the Royal Astronomical Society, 413*, L91–L95.
Volkenstein, M. (2009). *Entropy and information*. Basel: Birkhauser Physics.
Worrall, J. (1989). Structural realism: The best of both worlds. *Dialectica, 43*, 99–124.
Wray, B. K. (2014). Review of explaining science's success. *Australasian Journal of Philosophy*.
Wright, J. (1989). Realism and equivalence. *Erkenntnis, 31*, 109–128.
Wright, J. (1991). *Science and the theory of rationality*. Aldershot: Avebury.
Wright, J. (2013). *Explaining science's success: Understanding how scientific knowledge works*. Durham: Acumen Publishing Company.
Zalabardo, J. (2006). BonJour, externalism and the regress problem. *Synthese, 148*(1), 135–169.